MMK_dataset_analysis

Jianqiu Zheng

5/28/2020

dataset prep

This section of code extract unique soil samples into a new dataframe

```
library(ggplot2)
library(reshape2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
library(GGally)
```

```
##
## Attaching package: 'GGally'
```

```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

```
library(scales)
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
mpdata <- read.csv(file = 'datasets.csv') # dataset use a dummy colomn to scale all ra m
easurements
#==subsetting unique sample ====
rate<-dplyr::select(filter(mpdata,ra_s1==1), c(id, rs, mvol, mwp, bd, porosity, clay, or
g))

#===calculating accessible soc and oxy at optimum moisture condition (when ra==1)
opt<-function(po,mvol,bd,org){
  soc<-org/12*bd*1e-3  #assum ss 1% as DOC  (mol/cm3) and 10% of these DOC are accessibl
e
  oxy<-(po-mvol)*0.21*1.3e-6  #(mol/cm3)Henry's law constant (Caq/p=1.3e-3 mol/L/atm)
  return (list(soc, oxy))}
po<-rate$porosity
mvol<-rate$mvol
org<-rate$org
bd<-rate$bd
ss_opt<-opt(po,mvol,bd,org)
rate$soc_ss<-ss_opt[[1]] ##approximately 100mol/cm3
rate$oxy_ss<-ss_opt[[2]]

summary(rate)
```

```
##        id           rs              mvol             mwp
##  Ac     : 1   Min.   :0.1700   Min.   :0.1500   Min.   : 1.00
##  Ah     : 1   1st Qu.:0.4500   1st Qu.:0.2500   1st Qu.: 3.00
##  Bro1   : 1   Median :0.5500   Median :0.3000   Median : 6.00
##  Bro1c  : 1   Mean   :0.5433   Mean   :0.3072   Mean   :12.64
##  Bro2   : 1   3rd Qu.:0.6500   3rd Qu.:0.3600   3rd Qu.:14.95
##  Bro2c  : 1   Max.   :0.8200   Max.   :0.4870   Max.   :83.90
##  (Other):57
##        bd           porosity           clay             org
##  Min.   :0.460   Min.   :0.4300   Min.   :0.0300   Min.   :0.00700
##  1st Qu.:1.062   1st Qu.:0.5200   1st Qu.:0.1110   1st Qu.:0.01300
##  Median :1.150   Median :0.5600   Median :0.1750   Median :0.01600
##  Mean   :1.137   Mean   :0.5755   Mean   :0.1959   Mean   :0.02338
##  3rd Qu.:1.274   3rd Qu.:0.6065   3rd Qu.:0.2500   3rd Qu.:0.02100
##  Max.   :1.560   Max.   :0.8660   Max.   :0.5800   Max.   :0.12600
##
##      soc_ss              oxy_ss
##  Min.   :7.000e-07   Min.   :2.348e-08
##  1st Qu.:1.351e-06   1st Qu.:5.187e-08
##  Median :1.567e-06   Median :6.552e-08
##  Mean   :1.893e-06   Mean   :7.325e-08
##  3rd Qu.:1.926e-06   3rd Qu.:8.190e-08
##  Max.   :5.892e-06   Max.   :1.955e-07
##
```

Effective Km calculation

This section prepares parameters for effective Km calculations (based on Tang et al. 2019)

```
Npsite=3000 # number of transporter per cell
k2_p=100  #transporter specific substrate uptake rate, unit:s-1
rc=1.e-6 #microbial cell radius unit:m
rp=1.e-9  #transporter radius unit:m
Na=6.e23  #Avogadro number
Ratm=50 #atmospheric resistance, 50 s/m

temp=25+273.15#define temperature

#=calculating gaseous and aqueous tortuosity (Original paper based on Morldrup, 2003)
tau<-function (mvol, po, clay){
  wpo<-mvol  #water filled porosity
  gpo<-po-mvol #gas filled porosity
  b<-2.91+0.195*clay #shape parameter
  taug<- (po-mvol)*((po-mvol)/po)^3/b
  tauw<- mvol*(mvol/po)^(3/b-1)
  return (list(wpo=wpo, gpo=gpo,taug=taug,tauw=tauw))
}
mvol<-rate$mvol
po<-rate$porosity
clay<-rate$clay
tortuo<-tau(mvol, po,clay)

#==calculating O2 and C substrate diffusivity
Diffu<-function (gpo, wpo, taug, tauw){
  Dw_o2=1.4e-9*temp/298.0  #aqueous tracer diffusivity at 25
  Dg_o2=1.8e-5*(temp/273.0)**1.82 #oxygen diffusivity in gas phase
  henry_o2=3.2e-2*exp(-1500.*(1/temp-1/298.15))
  Dwo2= 0.5*(Dg_o2*taug*gpo*henry_o2+Dw_o2*tauw*wpo) #bulk aqueous molecular diffusivity
as a colume weighted average between aquesous and gaseous phases
  Dw_s=6e-9 #oxygen diffusivity in water
  Dws=Dw_s*tauw*wpo#bulk substrate diffusivity (between the soil matrix and microsite)
  return(list(Dwo2=Dwo2, Dws=Dws))
}

gpo<-tortuo$gpo
wpo<-tortuo$wpo
taug<-tortuo$taug
tauw<-tortuo$tauw
Dw<-Diffu(gpo, wpo, taug, tauw)

#==calculating effective microbial substrate affinity for OC and O2 (Tang et al 2019)===
=====
Kaff<-function(Ncell, gpo, wpo, taug, tauw, Dws,Dwg,mp){
  Bdens<-Ncell/Na #free microsite microbial abundance mol/m3
  Rm<-rc*(40*Ncell)^(1/3)  #microsite radius
  vm<-pi*4/3*Rm^3  #microsite volume
  kw2<-k2_p*Npsite #(Npsite*rp+pi*rc) #cell specific uptake rate for OX, unit s-1
  Dw_s0 = 6e-9  #for claculating reference affinity
  Dw_g0 = 1.4e-9
  fin<-Npsite*rp/ (Npsite*rp+pi*rc)#interception probability (number of molecules that 1
mol cell will encounter and be able to intercept)
  ksw1<-4*pi*Dw_s0*rc*fin*Na   #substrate delivery parameter unit m3 mo-1 s-1
```

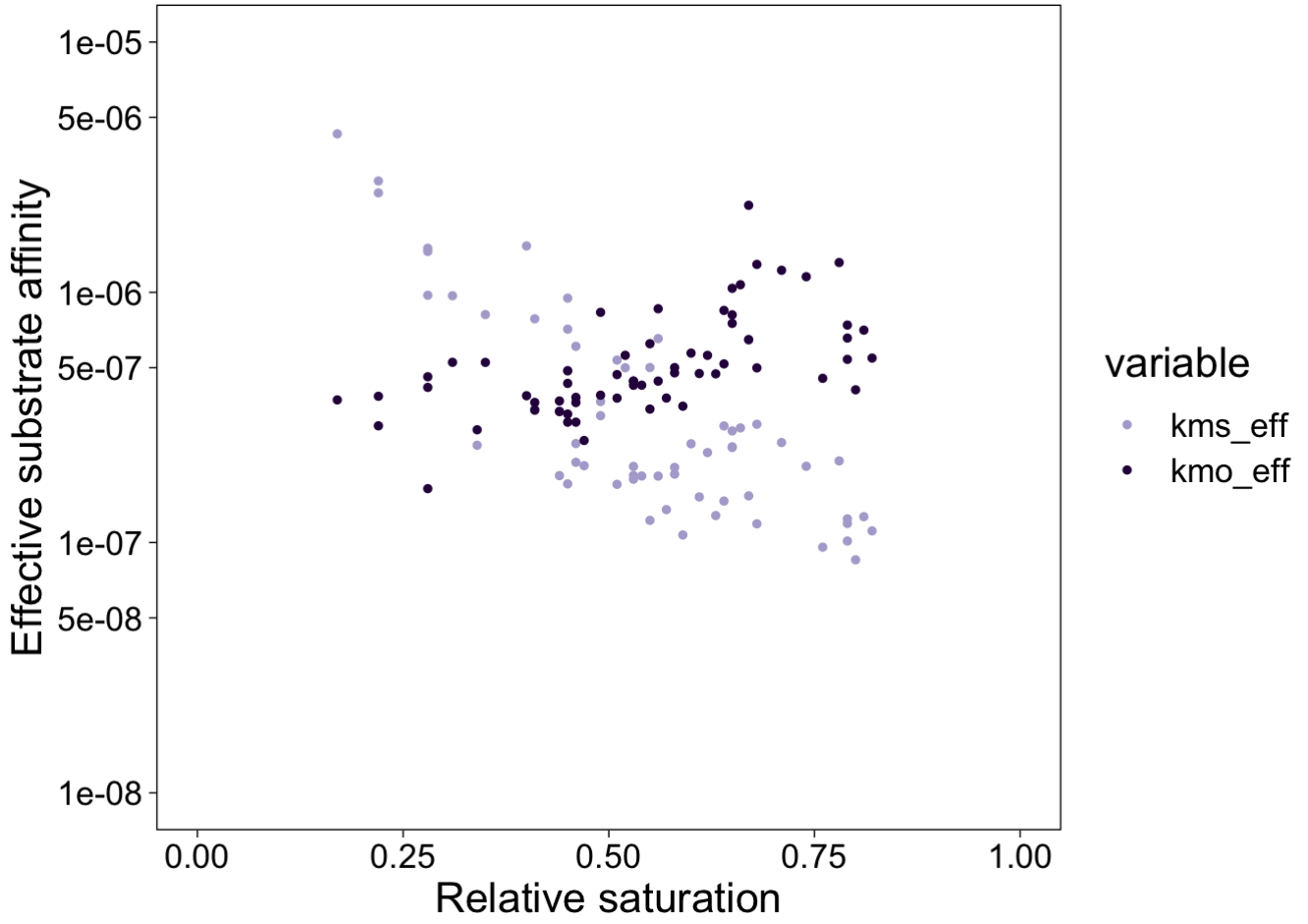
```
  kow1<-4*pi*Dw_g0*rc*fin*Na
  ksw0<-kw2/ksw1#reference affinity (Km used in MM kinetics in a well-mixed solution)
  kow0<-kw2/kow1#reference affinity (Km used in MM kinetics in a well-mixed solution)
  film<-exp(-13.65-0.857*log(mp/1000))    #calculting water film thickness
  ks_con<- (film/(Rm*Dw_s0*(Rm+film)) + 1/(Dws*(Rm+film)))*vm/(4*pi)#conductance coeffic
ient
  ks_aff<-ksw0*(1+ks_con*ksw1*Ncell/Na/vm)
  ko_con<- (film/(Rm*Dw_g0*(Rm+film)) + 1/(Dwg*(Rm+film)))*vm/(4*pi)#conductance coeffic
ient
  ko_aff<-kow0*(1+ko_con*kow1*Ncell/Na/vm)
  return(list(ksw0=ksw0, kow0=kow0,KaffOC=ks_aff, Kaffoxy=ko_aff))
}

mp<-rate$mwp
Dws<-Dw$Dws
Dwg<-Dw$Dwo2
Ncell<-2.68e-10*rate$soc_ss*1e-6*6.02e23
test<-Kaff(Ncell, gpo, wpo, taug, tauw, Dws,Dwg,mp)

rate$Keffs<-test$KaffOC*1e-3
rate$Keffo<-test$Kaffoxy*1e-3

summary(rate)
```

```
##        id           rs               mvol             mwp
##  Ac     : 1   Min.   :0.1700   Min.   :0.1500   Min.   : 1.00
##  Ah     : 1   1st Qu.:0.4500   1st Qu.:0.2500   1st Qu.: 3.00
##  Bro1   : 1   Median :0.5500   Median :0.3000   Median : 6.00
##  Bro1c  : 1   Mean   :0.5433   Mean   :0.3072   Mean   :12.64
##  Bro2   : 1   3rd Qu.:0.6500   3rd Qu.:0.3600   3rd Qu.:14.95
##  Bro2c  : 1   Max.   :0.8200   Max.   :0.4870   Max.   :83.90
##  (Other):57
##        bd           porosity           clay             org
##  Min.   :0.460   Min.   :0.4300   Min.   :0.0300   Min.   :0.00700
##  1st Qu.:1.062   1st Qu.:0.5200   1st Qu.:0.1110   1st Qu.:0.01300
##  Median :1.150   Median :0.5600   Median :0.1750   Median :0.01600
##  Mean   :1.137   Mean   :0.5755   Mean   :0.1959   Mean   :0.02338
##  3rd Qu.:1.274   3rd Qu.:0.6065   3rd Qu.:0.2500   3rd Qu.:0.02100
##  Max.   :1.560   Max.   :0.8660   Max.   :0.5800   Max.   :0.12600
##
##      soc_ss              oxy_ss              Keffs
##  Min.   :7.000e-07   Min.   :2.348e-08   Min.   :8.528e-08
##  1st Qu.:1.351e-06   1st Qu.:5.187e-08   1st Qu.:1.712e-07
##  Median :1.567e-06   Median :6.552e-08   Median :2.414e-07
##  Mean   :1.893e-06   Mean   :7.325e-08   Mean   :5.068e-07
##  3rd Qu.:1.926e-06   3rd Qu.:8.190e-08   3rd Qu.:5.183e-07
##  Max.   :5.892e-06   Max.   :1.955e-07   Max.   :4.298e-06
##
##      Keffo
##  Min.   :1.642e-07
##  1st Qu.:3.779e-07
##  Median :4.690e-07
##  Mean   :5.701e-07
##  3rd Qu.:6.355e-07
##  Max.   :2.227e-06
##
```

```
kmdata<-data.frame(rs=rate$rs)
kmdata$kms_eff<-rate$Keffs
kmdata$kmo_eff<-rate$Keffo
ptest<-reshape2::melt(kmdata, id.vars=("rs"))

base_breaks <- function(n = 10){
    function(x) {
        axisTicks(log10(range(x, na.rm = TRUE)), log = TRUE, n = n)
    }
}

plot1<-ggplot(ptest, aes(x = rs,y=value, color=variable)) +
  scale_x_continuous(name = "Relative saturation",limits = c(0,1)) +
  scale_y_continuous(name = "Effective substrate affinity",trans = log_trans(), breaks =
base_breaks(),limits = c(1e-8,1e-5))+
  geom_point(aes(color=variable,linetype=variable), size=1)+
  scale_color_manual(values = c("#b2abd2","#2d004b"))+
  scale_linetype_manual(values=c("solid","solid"))
```

```
## Warning: Ignoring unknown aesthetics: linetype
```

```
plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```



new dataframe for simulations

This section put the calculated accessible soc/oxy and effective affinity parameter Keffs/Keffo back to the orginal data frame.

```
userate<-dplyr::select(rate, c(id, soc_ss, oxy_ss,Keffs, Keffo))
usesoil<-dplyr::select(mpdata, c(id,rs,ra, mvol, mwp,  bd, porosity, clay, org))
newdata<-merge(usesoil, userate, by="id")
summary(newdata)
```

```
##        id            rs                ra               mvol
##  Ah     : 19   Min.   :0.0500   Min.   :0.0900   Min.   :0.0310
##  Ac     : 18   1st Qu.:0.2500   1st Qu.:0.6225   1st Qu.:0.1500
##  Bro1   :  8   Median :0.3900   Median :0.7900   Median :0.2360
##  Bro1c  :  8   Mean   :0.4188   Mean   :0.7385   Mean   :0.2399
##  Bro2   :  8   3rd Qu.:0.5800   3rd Qu.:0.9000   3rd Qu.:0.3297
##  Bro2c  :  8   Max.   :0.9000   Max.   :1.0000   Max.   :0.5150
##  (Other):349
##       mwp                bd           porosity           clay
##  Min.   :    1.0   Min.   :0.460   Min.   :0.3900   Min.   :0.0300
##  1st Qu.:    5.0   1st Qu.:1.040   1st Qu.:0.5300   1st Qu.:0.1200
##  Median :   23.0   Median :1.140   Median :0.5700   Median :0.1900
##  Mean   :  688.5   Mean   :1.111   Mean   :0.5817   Mean   :0.2036
##  3rd Qu.:  249.8   3rd Qu.:1.240   3rd Qu.:0.6082   3rd Qu.:0.2550
##  Max.   :13322.5   Max.   :1.597   Max.   :0.8660   Max.   :0.5800
##
##       org              soc_ss              oxy_ss
##  Min.   :0.00700   Min.   :7.000e-07   Min.   :2.348e-08
##  1st Qu.:0.01400   1st Qu.:1.370e-06   1st Qu.:4.750e-08
##  Median :0.01600   Median :1.585e-06   Median :6.306e-08
##  Mean   :0.02475   Mean   :1.968e-06   Mean   :7.216e-08
##  3rd Qu.:0.02200   3rd Qu.:1.927e-06   3rd Qu.:8.463e-08
##  Max.   :0.12600   Max.   :5.892e-06   Max.   :1.955e-07
##
##      Keffs               Keffo
##  Min.   :8.528e-08   Min.   :1.642e-07
##  1st Qu.:1.579e-07   1st Qu.:3.780e-07
##  Median :2.448e-07   Median :4.751e-07
##  Mean   :4.677e-07   Mean   :5.848e-07
##  3rd Qu.:4.995e-07   3rd Qu.:7.060e-07
##  Max.   :4.298e-06   Max.   :2.227e-06
##
```

##Diffusion based moisture-respiration relationship This section calculates mass transfer coefficients and analytical solution of diffusion-limited Michaelis-Menten microbial uptake kinetics

```
#==========Diffusion based moisture-respiration relationship I=========================
=====================================================
fh<-function(mvol,po,bd){
  Ds0=1.4e-9 #aqueous tracer diffusivity at 25 (m2/s)
  Dg0=2.1e-5 #oxygen diffusivity in gas
  fDg<-((po-mvol)/(po))^0.5#(po)^1.5*((po-mvol)/(po))^2.5 #gas phase relative diffusivit
y
  fDs<-((mvol)/(po))^0.5#(po)^1.5*((mvol)/(po))^2.5 #aqueous phase relative diffusivity
  H_o2<-1.3e-6  #mol/cm3/atm
  hs<-6/(mvol+bd*10)*Ds0*fDs/(0.00002^2) #DOM delivery (mass transfer rate in d-1)
  hg<-6/(mvol+bd*1)*Dg0*fDg*H_o2/(0.00002^2)  #DO delivery (mass transfer rate)
  return (list(fDs,fDg, hs, hg))}
mvol<-newdata$mvol
po<-newdata$porosity
bd<-newdata$bd
fhout<-fh(mvol, po, bd)
hs<-fhout[[3]]
hg<-fhout[[4]]

DiffMM<-function (hs,hg, soc, oxy, kmc, kmg,vmax){
  fm<-newdata$rs^(1/1.8)  #microbial hydrological sensitivity
  ac<-kmc/(soc)
  bc<-vmax/hs/(soc)
  t1c<-(1-4*bc/(1+ac+bc)^2)^0.5
  F1c<-(1+ac+bc)/2/bc*(1-t1c)
  css<-F1c*kmc/(1-F1c)
  ag<-kmg/(oxy)
  bg<-vmax/hg/(oxy)
  t1g<-(1-4*bg/(1+ag+bg)^2)^0.5
  F1g<-(1+ag+bg)/2/bg*(1-t1g)
  oss<-F1g*kmg/(1-F1g)
  Ft<-fm*F1c*F1g
  return(list(Ft))
}


soc<-newdata$soc_ss
oxy<-newdata$oxy_ss
vmax<-newdata$Vmax
kmo<-newdata$Keffo
kms<-newdata$Keffs

out1<-DiffMM(hs,hg, soc,oxy,1e-7,1e-7,1e-4) #Km_ref
out2<-DiffMM(hs,hg, soc,oxy,kms,kmo,1e-4) #Km_eff
newdata$out1<-out1[[1]]
newdata$out2<-out2[[1]]
newdata$out3<-newdata$rs^(1/2) *fhout[[1]]*fhout[[2]] #linear
newdata$out4<-3.11*newdata$rs-2.42*newdata$rs^2 #empirical
```

Rescale simulations

Rescale simulation results to 0-1 scale (because vmaxs are not optimized to ra measurements)

```
mpdata<-newdata
colnames(mpdata)[14]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out1_s<-new_scaled$scaled_new

mpdata<-newdata
colnames(mpdata)[15]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out2_s<-new_scaled$scaled_new


mpdata<-newdata
colnames(mpdata)[16]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

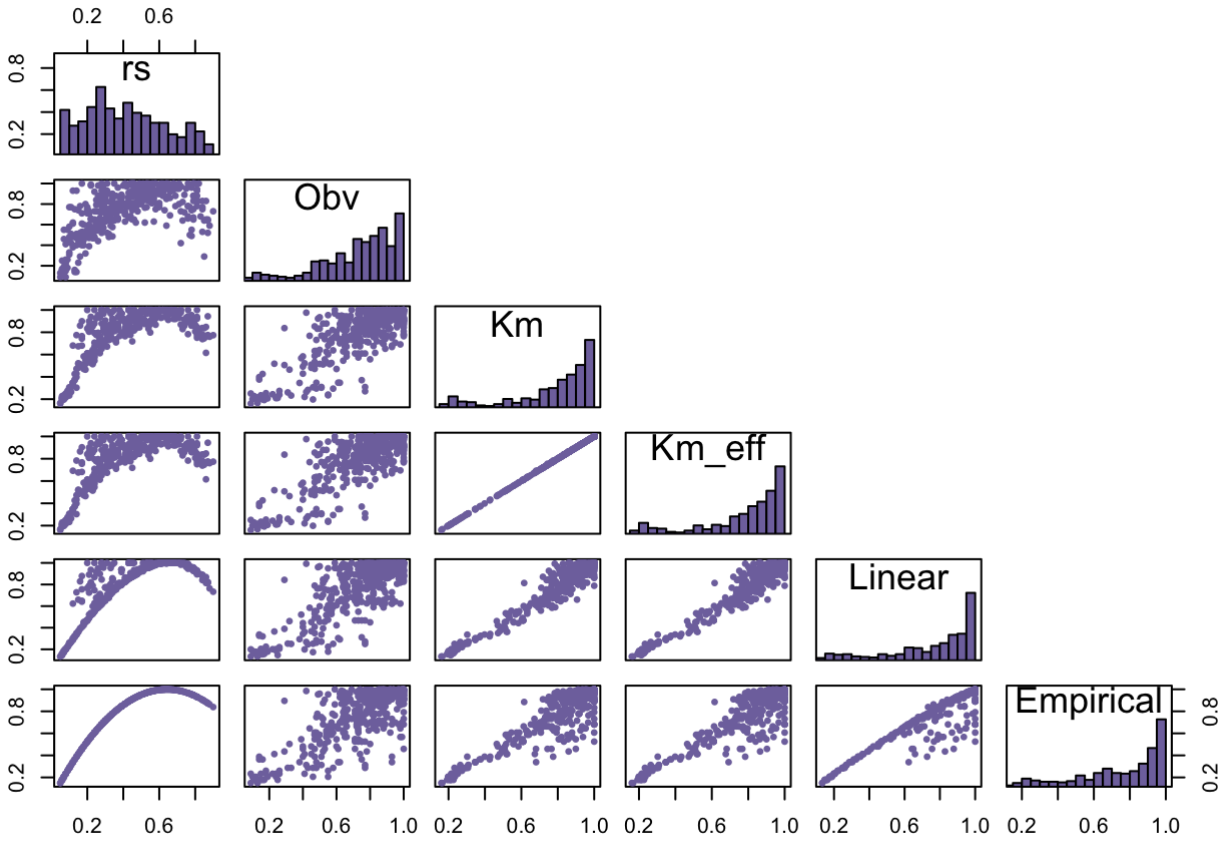
```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out3_s<-new_scaled$scaled_new
```

```
testplot<-data.frame(rs=newdata$rs)
testplot$Obv<-newdata$ra
testplot$Km<-newdata$out1_s
testplot$Km_eff<-newdata$out2_s
testplot$Linear<-newdata$out3_s
testplot$Empirical<-newdata$out4

#create matrix plot
panel.hist <- function(x, ...) {
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(usr[1:2], 0, 1.5) )
  h <- hist(x, breaks=20,plot = FALSE)
  breaks <- h$breaks; nB <- length(breaks)
  y <- h$counts; y <- y/max(y)
  rect(breaks[-nB], 0, breaks[-1], y, ...)
}


pairs(testplot, pch = 19,  cex = 0.5, col=c("#8073ac"),diag.panel=panel.hist,
      upper.panel=NULL)
```




```
dev.print(pdf,        # copies the plot to a the PDF file
          "Corr_matrix.pdf",width=7, height=5.8)
```

```
## quartz_off_screen
##                 2
```

##result statistics and visualization

```
ptest<-data.frame(rs=newdata$rs)
ptest$res<-newdata$out3_s-mpdata$ra

ave<-mean(mpdata$ra)
ssres<-sum(ptest$res^2)
sstot<-sum((mpdata$ra-ave)^2)
RR<-1-ssres/sstot

paste0("R-square = ", RR) #print R-square value
```

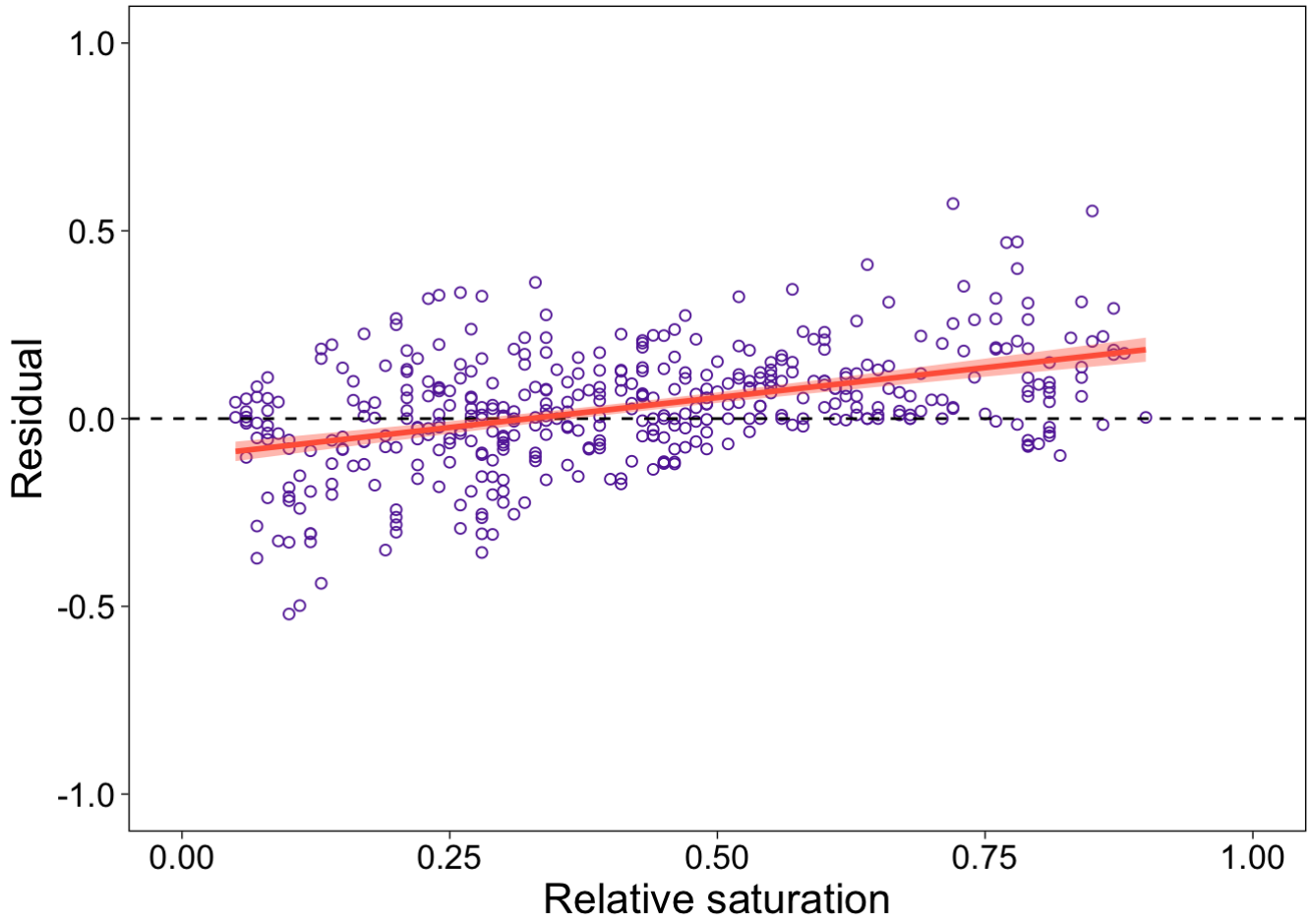
```
## [1] "R-square = 0.485235502885373"
```

```
simple.fit = lm(res~rs, data=ptest)
summary(simple.fit)
```

```
##
## Call:
## lm(formula = res ~ rs, data = ptest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44915 -0.08930  0.00236  0.08154  0.44621
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.10321    0.01452  -7.106 5.23e-12 ***
## rs           0.31860    0.03074  10.363  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1375 on 416 degrees of freedom
## Multiple R-squared:  0.2052, Adjusted R-squared:  0.2033
## F-statistic: 107.4 on 1 and 416 DF,  p-value: < 2.2e-16
```

```
#residual plot
plot1<-ggplot(ptest, aes(x = rs,y = res)) +
  scale_x_continuous(name = expression(paste("Relative saturation")),limits = c(0,1)) +
  scale_y_continuous(name = 'Residual',limits = c(-1,1))+geom_point(col="#642ba6", pch=1
, cex=1.6)+geom_hline(yintercept=0, color="black", linetype="dashed",size=0.5)+
  geom_smooth(method=lm, color="tomato1", size=1, fill="tomato1", se=TRUE)

plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Alternative diffusivity formulation

Bruggeman expression (Tjaden et al., 2016) Hamomoto expression (Hamamoto et al., 2010)

```
De<-function(mvol,po){
  Ds0=1.4e-9 #aqueous tracer diffusivity at 25 (m2/s)
  Dg0=2.1e-5 #oxygen diffusivity in gas
  DgB<-((po-mvol)/(po))^0.5 #Bruggeman expression
  DsB<-((mvol)/(po))^0.5
  DgH<-(po)^1.5*((po-mvol)/(po))^2.5 #Hamamoto expression
  DsH<-(po)^1.5*((mvol)/(po))^2.5
  return (list(DgB, DsB, DgH, DsH))}
mvol<-newdata$mvol
po<-newdata$porosity
Dout<-De(mvol, po)
DgB<-Dout[[1]]
DsB<-Dout[[2]]
DgH<-Dout[[3]]
DsH<-Dout[[4]]

SimLin<-function(mvol, po, fDs, fDg, soc, oxy){
  fm<-newdata$rs^(1/1.6)
  Ds0=1.4e-9 #aqueous tracer diffusivity at 25 (m2/s)
  Dg0=2.1e-5 #oxygen diffusivity in gas
  H_o2<-1.3e-6  #mol/cm3/atm
  hs<-6/(mvol+bd*10)*Ds0*fDs/(0.00002^2) #DOM delivery (mass transfer rate in d-1)
  hg<-6/(mvol+bd*1)*Dg0*fDg*H_o2/(0.00002^2)  #DO delivery (mass transfer rate)
  ft<-fm*hs*hg
  return (list(ft))}
soc<-newdata$soc_ss
oxy<-newdata$oxy_ss
mvol<-newdata$mvol
bd<-newdata$bd
po<-newdata$porosity
out1<-SimLin(mvol, po, DsB, DgB, soc,oxy)
out2<-SimLin(mvol, po, DsH, DgH, soc,oxy)
newdata$out1<-out1[[1]]
newdata$out2<-out2[[1]]


mpdata<-newdata
colnames(mpdata)[14]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out1_s<-new_scaled$scaled_new

mpdata<-newdata
colnames(mpdata)[15]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

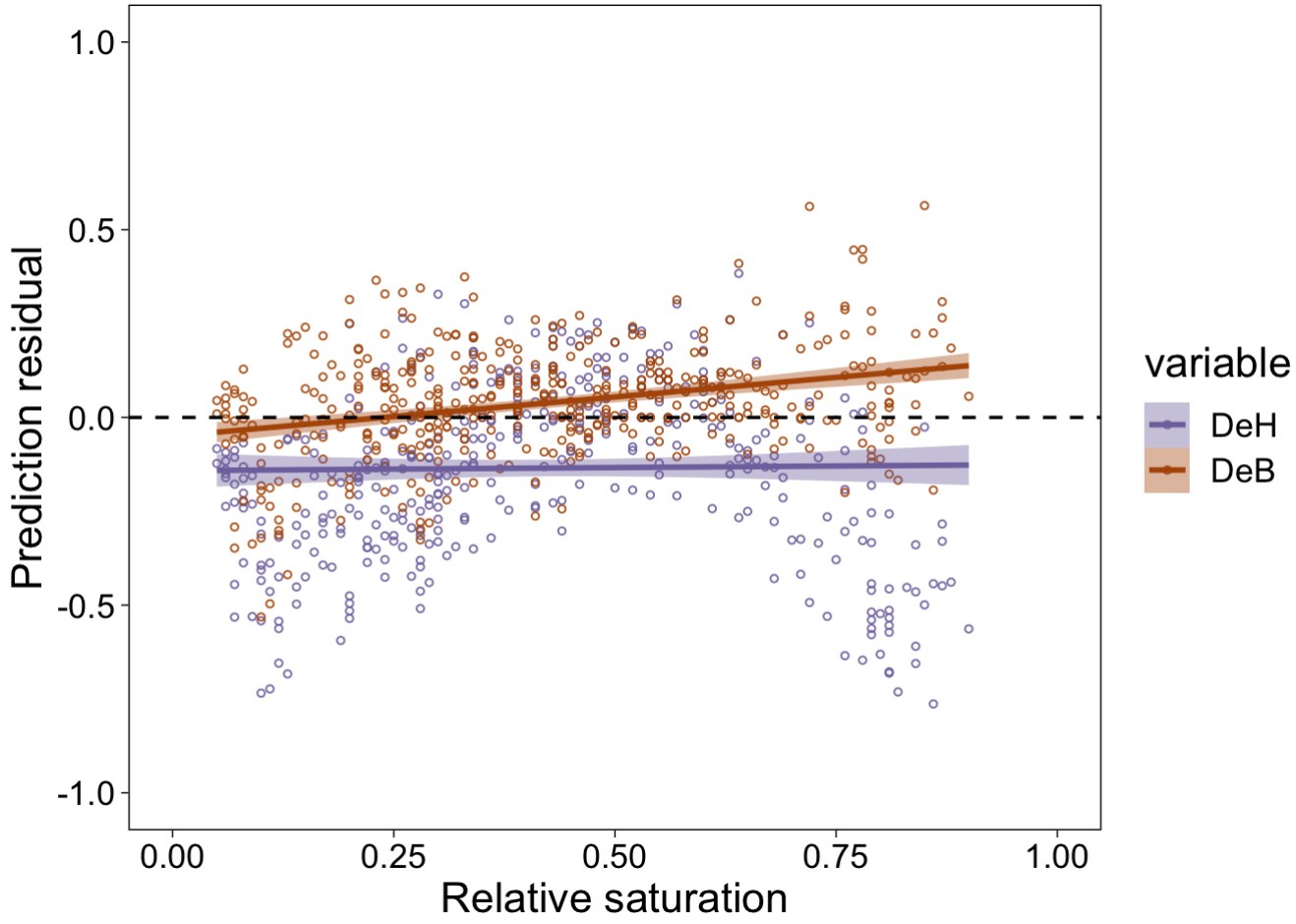
```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out2_s<-new_scaled$scaled_new

#===Residual plot====
diff<-data.frame(rs=newdata$rs)
diff$DeH<-newdata$out2_s-mpdata$ra
diff$DeB<-newdata$out1_s-mpdata$ra
pdiff<-reshape2::melt(diff, id.vars=("rs"))

plot1<-ggplot(pdiff, aes(x=rs, y=value))+geom_point(aes(color=variable),cex=1,shape=1, a
lpha=0.9)+
    geom_smooth(aes(color=variable, fill=variable),method = lm, se = TRUE, size=1)+
    geom_hline(yintercept=0, color="black",linetype="dashed",size=0.6)+
    scale_color_manual(values = c("#8073ac","#b35806","#662788"))+
```

```
    scale_fill_manual(values = c("#8073ac","#b35806","#662788"))+
    scale_x_continuous(name = 'Relative saturation',limits = c(0,1))+
    scale_y_continuous(name = 'Prediction residual',limits = c(-1,1))+
    theme(text = element_text(size=16))
plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
pdf("trait_Diff.pdf", width=4.4, height=3)
plot2
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
dev.off()
```

```
## quartz_off_screen
##                 2
```

```
#====statistics====
ptest<-data.frame(rs=newdata$rs)
ptest$res<-newdata$out1_s-mpdata$ra

ave<-mean(mpdata$ra)
ssres<-sum(ptest$res^2)
sstot<-sum((mpdata$ra-ave)^2)
RR<-1-ssres/sstot

paste0("R-square = ", RR) #print R-square value
```

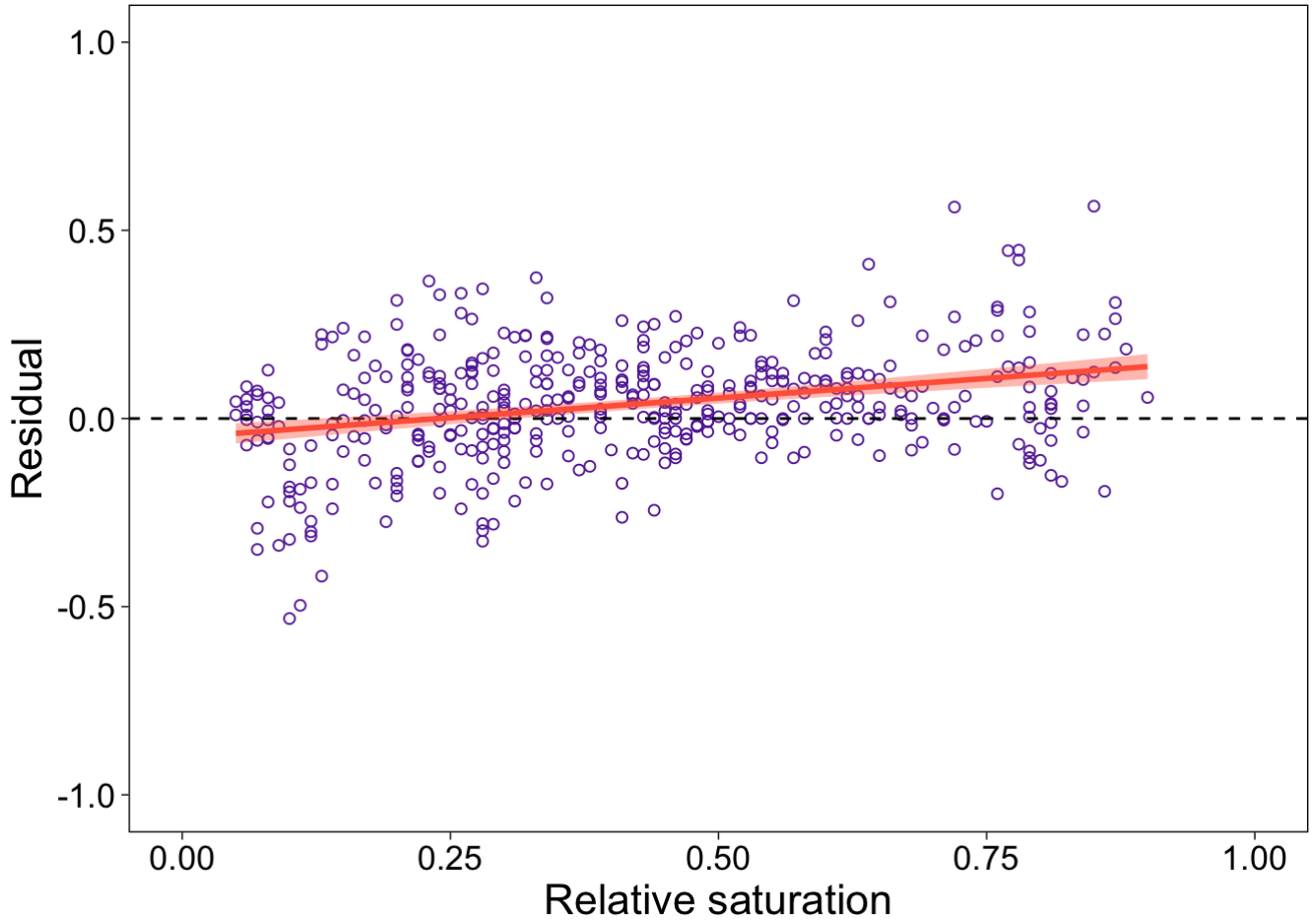
```
## [1] "R-square = 0.49655628697962"
```

```
simple.fit = lm(res~rs, data=ptest)
summary(simple.fit)
```

```
##
## Call:
## lm(formula = res ~ rs, data = ptest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50260 -0.08370  0.00050  0.08855  0.46153
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.04986    0.01518  -3.285  0.00111 **
## rs           0.20870    0.03213   6.496 2.37e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1437 on 416 degrees of freedom
## Multiple R-squared:  0.09208,    Adjusted R-squared:  0.0899
## F-statistic: 42.19 on 1 and 416 DF,  p-value: 2.366e-10
```

```
#residual plot
plot1<-ggplot(ptest, aes(x = rs,y = res)) +
  scale_x_continuous(name = expression(paste("Relative saturation")),limits = c(0,1)) +
  scale_y_continuous(name = 'Residual',limits = c(-1,1))+geom_point(col="#642ba6", pch=1
, cex=1.6)+geom_hline(yintercept=0, color="black", linetype="dashed",size=0.5)+
  geom_smooth(method=lm, color="tomato1", size=1, fill="tomato1", se=TRUE)

plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Alternative kinetic formulation

```
SimKin<-function(mvol, po, soc, oxy, kmc, kmg,vmax){
  fm<-newdata$rs^(1/1.6)
  Ds0=1.4e-9 #aqueous tracer diffusivity at 25 (m2/s)
  Dg0=2.1e-5 #oxygen diffusivity in gas
  fDg<-((po-mvol)/(po))^0.5 #gas phase relative diffusivity
  fDs<-((mvol)/(po))^0.5 #aqueous phase relative diffusivity
  H_o2<-1.3e-6  #mol/cm3/atm
  hs<-6/(mvol+bd*10)*Ds0*fDs/(0.00002^2) #DOM delivery (mass transfer rate in d-1)
  hg<-6/(mvol+bd*1)*Dg0*fDg*H_o2/(0.00002^2)  #DO delivery (mass transfer rate)
  ac<-kmc/(soc)
  bc<-vmax/hs/(soc)
  t1c<-(1-4*bc/(1+ac+bc)^2)^0.5
  F1c<-(1+ac+bc)/2/bc*(1-t1c)
  ag<-kmg/(oxy)
  bg<-vmax/hg/(oxy)
  t1g<-(1-4*bg/(1+ag+bg)^2)^0.5
  F1g<-(1+ag+bg)/2/bg*(1-t1g)
  ft1<-fm*vmax*F1c*F1g
  ft2<-fm*hs*hg
  return (list(ft1,ft2))}
soc<-newdata$soc_ss
oxy<-newdata$oxy_ss
mvol<-newdata$mvol
po<-newdata$porosity
out1<-SimKin(mvol, po, soc,oxy,1e-7,1e-7, 2.2e-6)
newdata$out1<-out1[[1]]
newdata$out2<-out1[[2]]

mpdata<-newdata
colnames(mpdata)[14]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out1_s<-new_scaled$scaled_new

mpdata<-newdata
colnames(mpdata)[15]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

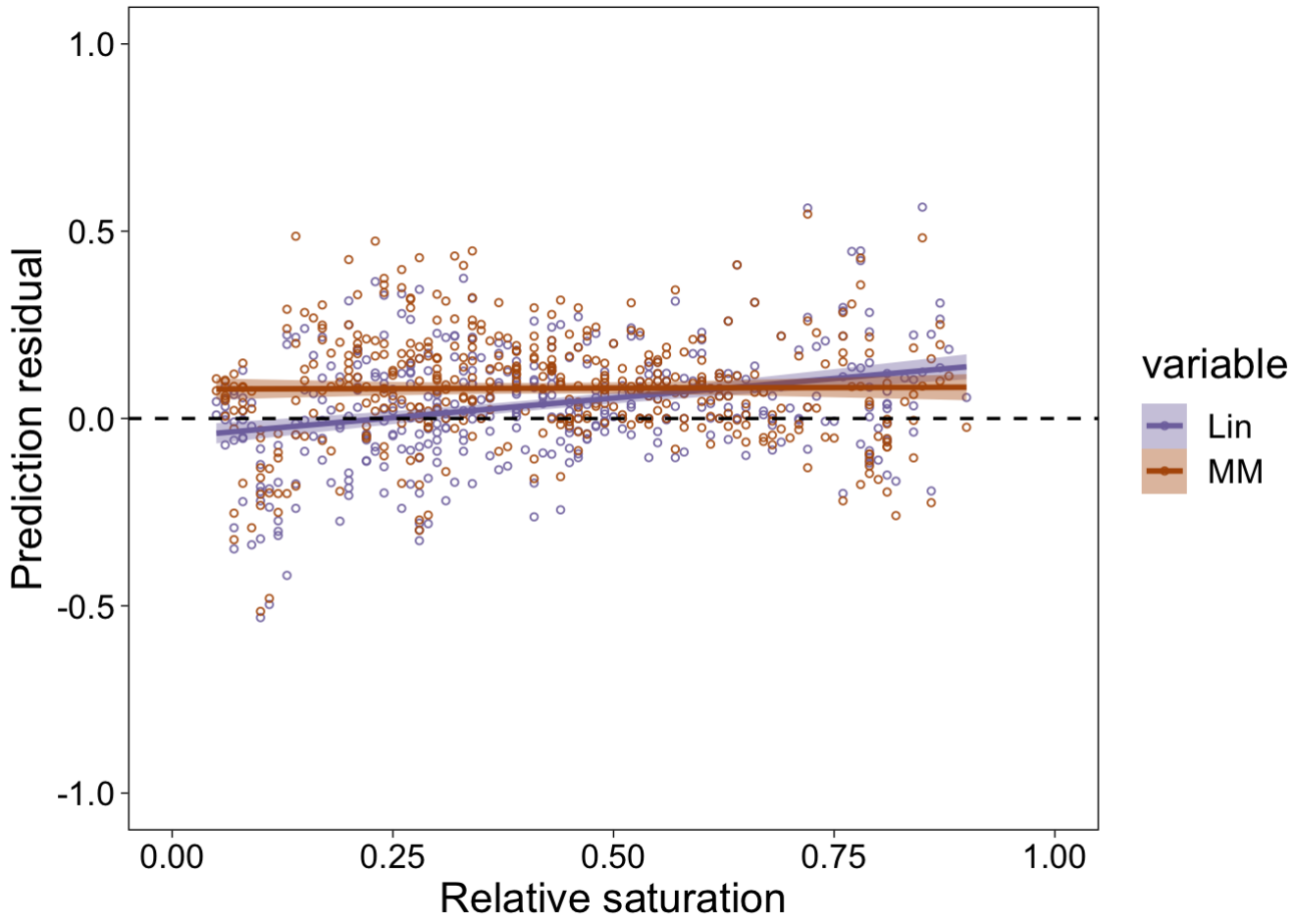
```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out2_s<-new_scaled$scaled_new

#===Residual plot====
diff<-data.frame(rs=newdata$rs)
diff$Lin<-newdata$out2_s-mpdata$ra
diff$MM<-newdata$out1_s-mpdata$ra
pdiff<-reshape2::melt(diff, id.vars=("rs"))

plot1<-ggplot(pdiff, aes(x=rs, y=value))+geom_point(aes(color=variable),cex=1,shape=1, a
lpha=0.9)+
    geom_smooth(aes(color=variable, fill=variable),method = lm, se = TRUE, size=1)+
    geom_hline(yintercept=0, color="black",linetype="dashed",size=0.6)+
    scale_color_manual(values = c("#8073ac","#b35806","#662788"))+
```

```
    scale_fill_manual(values = c("#8073ac","#b35806","#662788"))+
    scale_x_continuous(name = 'Relative saturation',limits = c(0,1))+
    scale_y_continuous(name = 'Prediction residual',limits = c(-1,1))+
    theme(text = element_text(size=16))
plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
pdf("trait_Kin.pdf", width=4.4, height=3)
plot2
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
dev.off()
```

```
## quartz_off_screen
##                 2
```

```
#====statistics====
ptest<-data.frame(rs=newdata$rs)
ptest$res<-newdata$out1_s-mpdata$ra

ave<-mean(mpdata$ra)
ssres<-sum(ptest$res^2)
sstot<-sum((mpdata$ra-ave)^2)
RR<-1-ssres/sstot

paste0("R-square = ", RR) #print R-square value
```

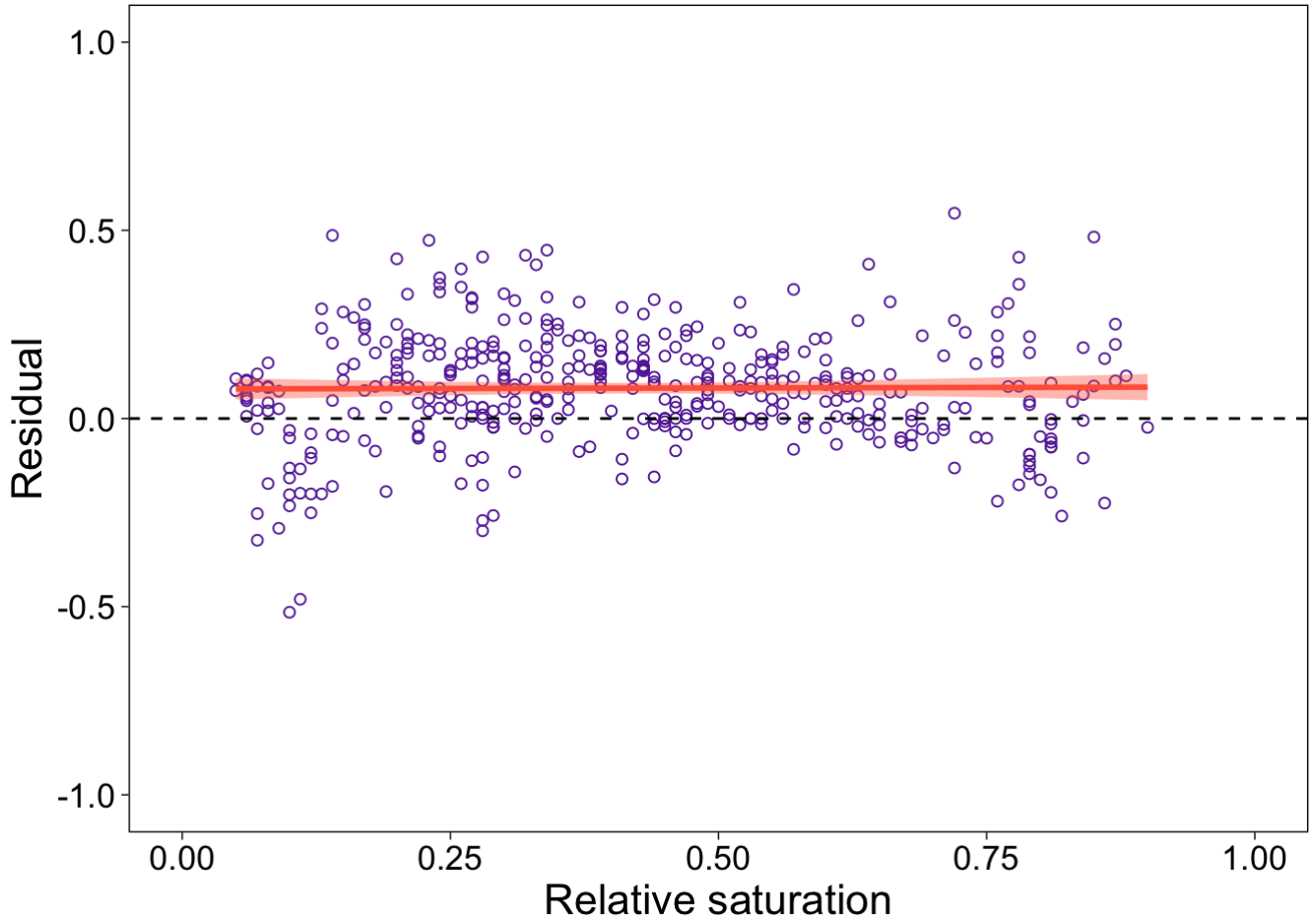
```
## [1] "R-square = 0.398257863736381"
```

```
simple.fit = lm(res~rs, data=ptest)
summary(simple.fit)
```

```
##
## Call:
## lm(formula = res ~ rs, data = ptest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.59391 -0.08185  0.00360  0.08821  0.46318
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.078468   0.015781   4.972 9.69e-07 ***
## rs          0.005534   0.033402   0.166    0.868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1494 on 416 degrees of freedom
## Multiple R-squared:  6.599e-05,  Adjusted R-squared:  -0.002338
## F-statistic: 0.02745 on 1 and 416 DF,  p-value: 0.8685
```

```
#residual plot
plot1<-ggplot(ptest, aes(x = rs,y = res)) +
  scale_x_continuous(name = expression(paste("Relative saturation")),limits = c(0,1)) +
  scale_y_continuous(name = 'Residual',limits = c(-1,1))+geom_point(col="#642ba6", pch=1
, cex=1.6)+geom_hline(yintercept=0, color="black", linetype="dashed",size=0.5)+
  geom_smooth(method=lm, color="tomato1", size=1, fill="tomato1", se=TRUE)

plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Alternative formulation of microbial hydrological sensitivity

```
SimKin<-function(mvol, po, soc, oxy, kmc, kmg,vmax){
  fm1<-newdata$rs^(1/1.6)
  fm2<-exp(-0.0004*newdata$mwp)
  Ds0=1.4e-9 #aqueous tracer diffusivity at 25 (m2/s)
  Dg0=2.1e-5 #oxygen diffusivity in gas
  fDg<-((po-mvol)/(po))^0.5 #gas phase relative diffusivity
  fDs<-((mvol)/(po))^0.5 #aqueous phase relative diffusivity
  H_o2<-1.3e-6  #mol/cm3/atm
  hs<-6/(mvol+bd*10)*Ds0*fDs/(0.00002^2) #DOM delivery (mass transfer rate in d-1)
  hg<-6/(mvol+bd*1)*Dg0*fDg*H_o2/(0.00002^2)  #DO delivery (mass transfer rate)
  ac<-kmc/(soc)
  bc<-vmax/hs/(soc)
  t1c<-(1-4*bc/(1+ac+bc)^2)^0.5
  F1c<-(1+ac+bc)/2/bc*(1-t1c)
  ag<-kmg/(oxy)
  bg<-vmax/hg/(oxy)
  t1g<-(1-4*bg/(1+ag+bg)^2)^0.5
  F1g<-(1+ag+bg)/2/bg*(1-t1g)
  ft1<-fm1*vmax*F1c*F1g
  ft2<-fm2*vmax*F1c*F1g
  return (list(ft1,ft2))}
soc<-newdata$soc_ss
oxy<-newdata$oxy_ss
mvol<-newdata$mvol
po<-newdata$porosity
out1<-SimKin(mvol, po, soc,oxy,1e-7,1e-7, 2.2e-6)
newdata$out1<-out1[[1]]
newdata$out2<-out1[[2]]

mpdata<-newdata
colnames(mpdata)[14]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out1_s<-new_scaled$scaled_new

mpdata<-newdata
colnames(mpdata)[15]<-("pred_new")
mpdata$scaled_new<-0

#check unique id======
unique(mpdata$id)
```

```
##  [1] Ac          Ah          Bro1        Bro1c       Bro2
##  [6] Bro2c       Bro3        Bro3c       Bro4        Bro4c
## [11] Bro5        Bro5c       Cecile      Clarion     CloMin1
## [16] CloMin1c    CloMin2     CloMin2c    CloMin3     CloMin3c
## [21] CloMin4     CloMin4c    CloMin5     CloMin5c    Crider
## [26] dry_btm     dry_top     FortCollins Frederick   Houston
## [31] Int_btm     Int_mid     Int_top     Kole        LERB_CL2
## [36] LERB_CL4    LERB_CL6    LERB_NA1    LERB_NA2    LERB_NA3
## [41] LERB_NA4    LERB_NA5    LERB_NA6    LERB_SI2    LERB_SI4
## [46] LERB_SI6    Miami       Min1        Min1c       Min2
## [51] Min2c       Min3        Min3c       Min4        Min4c
## [56] Min5        Min5c       Mohave      Valentine   Wahiawa
## [61] Walla       wet_btm     wet_mid
## 63 Levels: Ac Ah Bro1 Bro1c Bro2 Bro2c Bro3 Bro3c Bro4 Bro4c ... wet_mid
```

```
#id1-12 Ac/Ah/Bro1/Bro1c/Bro2/Bro2c/Bro3/Bro3c/Bro4 /Bro4c/Bro5/Bro5c
id_1<-subset(mpdata, id=="Ac")
id_1$scaled_new<-id_1$pred_new/max(id_1$pred_new)

id_2<-subset(mpdata, id=="Ah")
id_2$scaled_new<-id_2$pred_new/max(id_2$pred_new)

id_3<-subset(mpdata, id=="Bro1")
id_3$scaled_new<-id_3$pred_new/max(id_3$pred_new)

id_4<-subset(mpdata, id=="Bro1c")
id_4$scaled_new<-id_4$pred_new/max(id_4$pred_new)

id_5<-subset(mpdata, id=="Bro2")
id_5$scaled_new<-id_5$pred_new/max(id_5$pred_new)

id_6<-subset(mpdata, id=="Bro2c")
id_6$scaled_new<-id_6$pred_new/max(id_6$pred_new)

id_7<-subset(mpdata, id=="Bro3")
id_7$scaled_new<-id_7$pred_new/max(id_7$pred_new)

id_8<-subset(mpdata, id=="Bro3c")
id_8$scaled_new<-id_8$pred_new/max(id_8$pred_new)

id_9<-subset(mpdata, id=="Bro4")
id_9$scaled_new<-id_9$pred_new/max(id_9$pred_new)

id_10<-subset(mpdata, id=="Bro4c")
id_10$scaled_new<-id_10$pred_new/max(id_10$pred_new)

id_11<-subset(mpdata, id=="Bro5")
id_11$scaled_new<-id_11$pred_new/max(id_11$pred_new)

id_12<-subset(mpdata, id=="Bro5c")
id_12$scaled_new<-id_12$pred_new/max(id_12$pred_new)
#Cecile/Clarion/CloMin1/CloMin1c/CloMin2/CloMin2c/CloMin3 / CloMin3c /CloMin4 /CloMin4c/
CloMin5 / CloMin5c
id_13<-subset(mpdata, id=="Cecile")
id_13$scaled_new<-id_13$pred_new/max(id_13$pred_new)

id_14<-subset(mpdata, id=="Clarion")
id_14$scaled_new<-id_14$pred_new/max(id_14$pred_new)

id_15<-subset(mpdata, id=="CloMin1")
id_15$scaled_new<-id_15$pred_new/max(id_15$pred_new)

id_16<-subset(mpdata, id=="CloMin1c")
id_16$scaled_new<-id_16$pred_new/max(id_16$pred_new)

id_17<-subset(mpdata, id=="CloMin2")
id_17$scaled_new<-id_17$pred_new/max(id_17$pred_new)
```

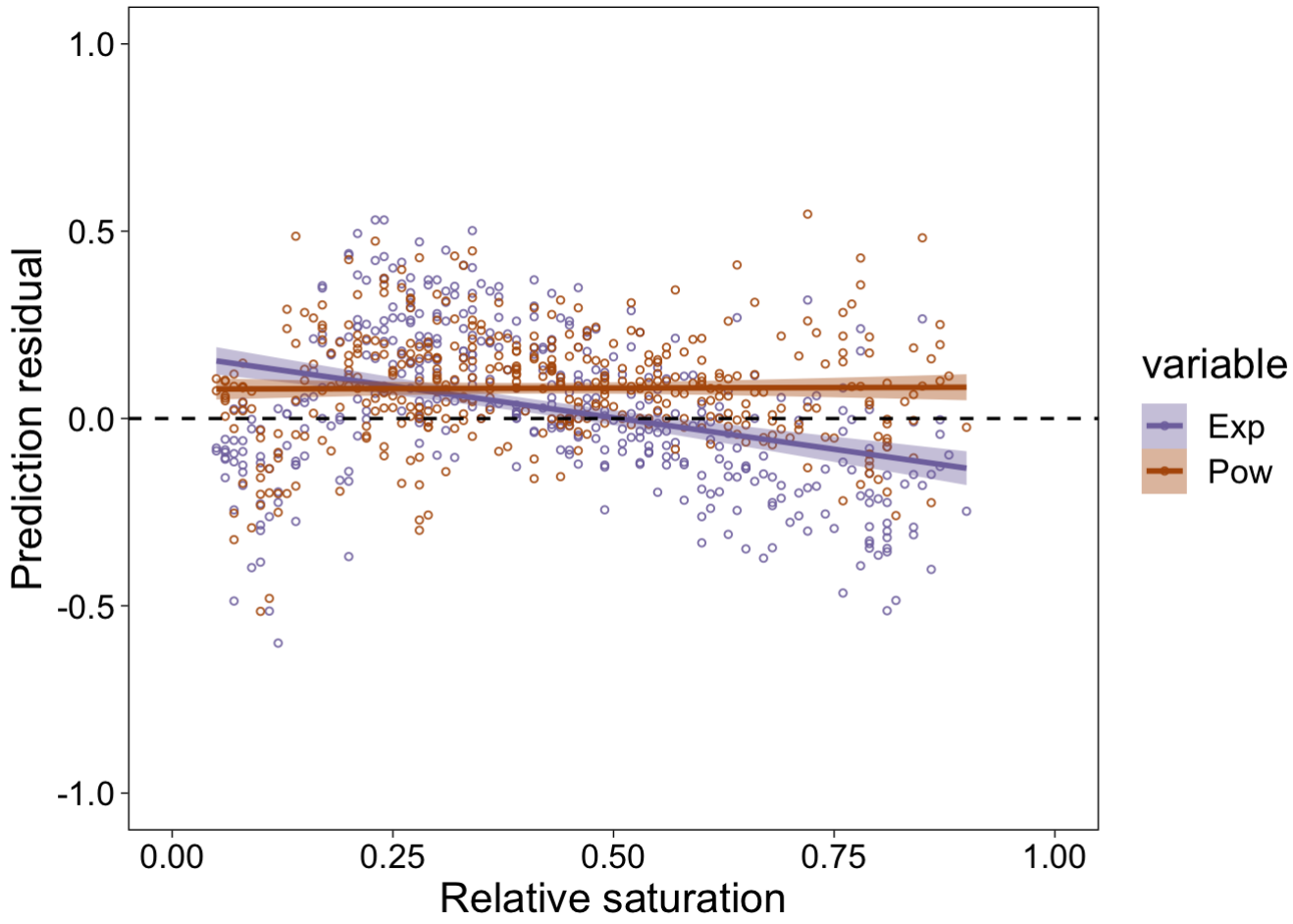
```
id_18<-subset(mpdata, id=="CloMin2c")
id_18$scaled_new<-id_18$pred_new/max(id_18$pred_new)

id_19<-subset(mpdata, id=="CloMin3")
id_19$scaled_new<-id_19$pred_new/max(id_19$pred_new)

id_20<-subset(mpdata, id=="CloMin3c")
id_20$scaled_new<-id_20$pred_new/max(id_20$pred_new)

id_21<-subset(mpdata, id=="CloMin4")
id_21$scaled_new<-id_21$pred_new/max(id_21$pred_new)

id_22<-subset(mpdata, id=="CloMin4c")
id_22$scaled_new<-id_22$pred_new/max(id_22$pred_new)

id_23<-subset(mpdata, id=="CloMin5")
id_23$scaled_new<-id_23$pred_new/max(id_23$pred_new)

id_24<-subset(mpdata, id=="CloMin5c")
id_24$scaled_new<-id_24$pred_new/max(id_24$pred_new)
#Crider/dry_btm /dry_top /FortCollins /Frederick /Houston / Int_btm  / Int_mid /Int_top/
Kole / LERB_CL2/LERB_CL4
id_25<-subset(mpdata, id=="Crider")
id_25$scaled_new<-id_25$pred_new/max(id_25$pred_new)

id_26<-subset(mpdata, id=="dry_btm")
id_26$scaled_new<-id_26$pred_new/max(id_26$pred_new)

id_27<-subset(mpdata, id=="dry_top")
id_27$scaled_new<-id_27$pred_new/max(id_27$pred_new)

id_28<-subset(mpdata, id=="FortCollins")
id_28$scaled_new<-id_28$pred_new/max(id_28$pred_new)

id_29<-subset(mpdata, id=="Frederick")
id_29$scaled_new<-id_29$pred_new/max(id_29$pred_new)

id_30<-subset(mpdata, id=="Houston")
id_30$scaled_new<-id_30$pred_new/max(id_30$pred_new)

id_31<-subset(mpdata, id=="Int_btm")
id_31$scaled_new<-id_31$pred_new/max(id_31$pred_new)

id_32<-subset(mpdata, id=="Int_mid")
id_32$scaled_new<-id_32$pred_new/max(id_32$pred_new)

id_33<-subset(mpdata, id=="Int_top")
id_33$scaled_new<-id_33$pred_new/max(id_33$pred_new)

id_34<-subset(mpdata, id=="Kole")
id_34$scaled_new<-id_34$pred_new/max(id_34$pred_new)

id_35<-subset(mpdata, id=="LERB_CL2")
id_35$scaled_new<-id_35$pred_new/max(id_35$pred_new)
```

```
id_36<-subset(mpdata, id=="LERB_CL4")
id_36$scaled_new<-id_36$pred_new/max(id_36$pred_new)
#LERB_CL6/LERB_NA1 /LERB_NA2 /LERB_NA3 /LERB_NA4 /LERB_NA5/LERB_NA6 /LERB_SI2/LERB_SI4 /
LERB_SI6/ Miami /Min1
id_37<-subset(mpdata, id=="LERB_CL6")
id_37$scaled_new<-id_37$pred_new/max(id_37$pred_new)

id_38<-subset(mpdata, id=="LERB_NA1")
id_38$scaled_new<-id_38$pred_new/max(id_38$pred_new)

id_39<-subset(mpdata, id=="LERB_NA2")
id_39$scaled_new<-id_39$pred_new/max(id_39$pred_new)

id_40<-subset(mpdata, id=="LERB_NA3")
id_40$scaled_new<-id_40$pred_new/max(id_40$pred_new)

id_41<-subset(mpdata, id=="LERB_NA4")
id_41$scaled_new<-id_41$pred_new/max(id_41$pred_new)

id_42<-subset(mpdata, id=="LERB_NA5")
id_42$scaled_new<-id_42$pred_new/max(id_42$pred_new)

id_43<-subset(mpdata, id=="LERB_NA6")
id_43$scaled_new<-id_43$pred_new/max(id_43$pred_new)

id_44<-subset(mpdata, id=="LERB_SI2")
id_44$scaled_new<-id_44$pred_new/max(id_44$pred_new)

id_45<-subset(mpdata, id=="LERB_SI4")
id_45$scaled_new<-id_45$pred_new/max(id_45$pred_new)

id_46<-subset(mpdata, id=="LERB_SI6")
id_46$scaled_new<-id_46$pred_new/max(id_46$pred_new)

id_47<-subset(mpdata, id=="Miami")
id_47$scaled_new<-id_47$pred_new/max(id_47$pred_new)

id_48<-subset(mpdata, id=="Min1")
id_48$scaled_new<-id_48$pred_new/max(id_48$pred_new)
#Min1c       Min2        Min2c       Min3        Min3c       Min4        Min4c       Min
5        Min5c       Mohave      Valentine   Wahiawa
id_49<-subset(mpdata, id=="Min1c")
id_49$scaled_new<-id_49$pred_new/max(id_49$pred_new)

id_50<-subset(mpdata, id=="Min2")
id_50$scaled_new<-id_50$pred_new/max(id_50$pred_new)

id_51<-subset(mpdata, id=="Min2c")
id_51$scaled_new<-id_51$pred_new/max(id_51$pred_new)

id_52<-subset(mpdata, id=="Min3")
id_52$scaled_new<-id_52$pred_new/max(id_52$pred_new)
```

```
id_53<-subset(mpdata, id=="Min3c")
id_53$scaled_new<-id_53$pred_new/max(id_53$pred_new)

id_54<-subset(mpdata, id=="Min4")
id_54$scaled_new<-id_54$pred_new/max(id_54$pred_new)

id_55<-subset(mpdata, id=="Min4c")
id_55$scaled_new<-id_55$pred_new/max(id_55$pred_new)

id_56<-subset(mpdata, id=="Min5")
id_56$scaled_new<-id_56$pred_new/max(id_56$pred_new)

id_57<-subset(mpdata, id=="Min5c")
id_57$scaled_new<-id_57$pred_new/max(id_57$pred_new)

id_58<-subset(mpdata, id=="Mohave")
id_58$scaled_new<-id_58$pred_new/max(id_58$pred_new)

id_59<-subset(mpdata, id=="Valentine")
id_59$scaled_new<-id_59$pred_new/max(id_59$pred_new)

id_60<-subset(mpdata, id=="Wahiawa")
id_60$scaled_new<-id_60$pred_new/max(id_60$pred_new)
#id61-63 Walla       wet_btm     wet_mid
id_61<-subset(mpdata, id=="Walla")
id_61$scaled_new<-id_61$pred_new/max(id_61$pred_new)

id_62<-subset(mpdata, id=="wet_btm")
id_62$scaled_new<-id_62$pred_new/max(id_62$pred_new)

id_63<-subset(mpdata, id=="wet_mid")
id_63$scaled_new<-id_63$pred_new/max(id_63$pred_new)

new_scaled<-rbind(id_1,id_2,id_3,id_4,id_5,id_6,id_7,id_8,id_9,id_10,id_11,id_12,id_13,i
d_14,id_15,id_16,id_17,id_18,id_19,id_20,
                  id_21,id_22,id_23,id_24,id_25,id_26,id_27,id_28,id_29,id_30,id_31,id_3
2,id_33,id_34,id_35,id_36,id_37,id_38,id_39,id_40,
                  id_41,id_42,id_43,id_44,id_45,id_46,id_47,id_48,id_49,id_50,id_51,id_5
2,id_53,id_54,id_55,id_56,id_57,id_58,id_59,id_60,
                  id_61,id_62,id_63)
#=====
newdata$out2_s<-new_scaled$scaled_new

#===Residual plot====
diff<-data.frame(rs=newdata$rs)
diff$Exp<-newdata$out2_s-mpdata$ra
diff$Pow<-newdata$out1_s-mpdata$ra
pdiff<-reshape2::melt(diff, id.vars=("rs"))

plot1<-ggplot(pdiff, aes(x=rs, y=value))+geom_point(aes(color=variable),cex=1,shape=1, a
lpha=0.9)+
    geom_smooth(aes(color=variable, fill=variable),method = lm, se = TRUE, size=1)+
    geom_hline(yintercept=0, color="black",linetype="dashed",size=0.6)+
    scale_color_manual(values = c("#8073ac","#b35806","#662788"))+
```

```
    scale_fill_manual(values = c("#8073ac","#b35806","#662788"))+
    scale_x_continuous(name = 'Relative saturation',limits = c(0,1))+
    scale_y_continuous(name = 'Prediction residual',limits = c(-1,1))+
    theme(text = element_text(size=16))
plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
pdf("trait_Phy.pdf", width=4.4, height=3)
plot2
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
dev.off()
```

```
## quartz_off_screen
##                 2
```

```
#====statistics====
ptest<-data.frame(rs=newdata$rs)
ptest$res<-newdata$out2_s-mpdata$ra

ave<-mean(mpdata$ra)
ssres<-sum(ptest$res^2)
sstot<-sum((mpdata$ra-ave)^2)
RR<-1-ssres/sstot

paste0("R-square = ", RR) #print R-square value
```

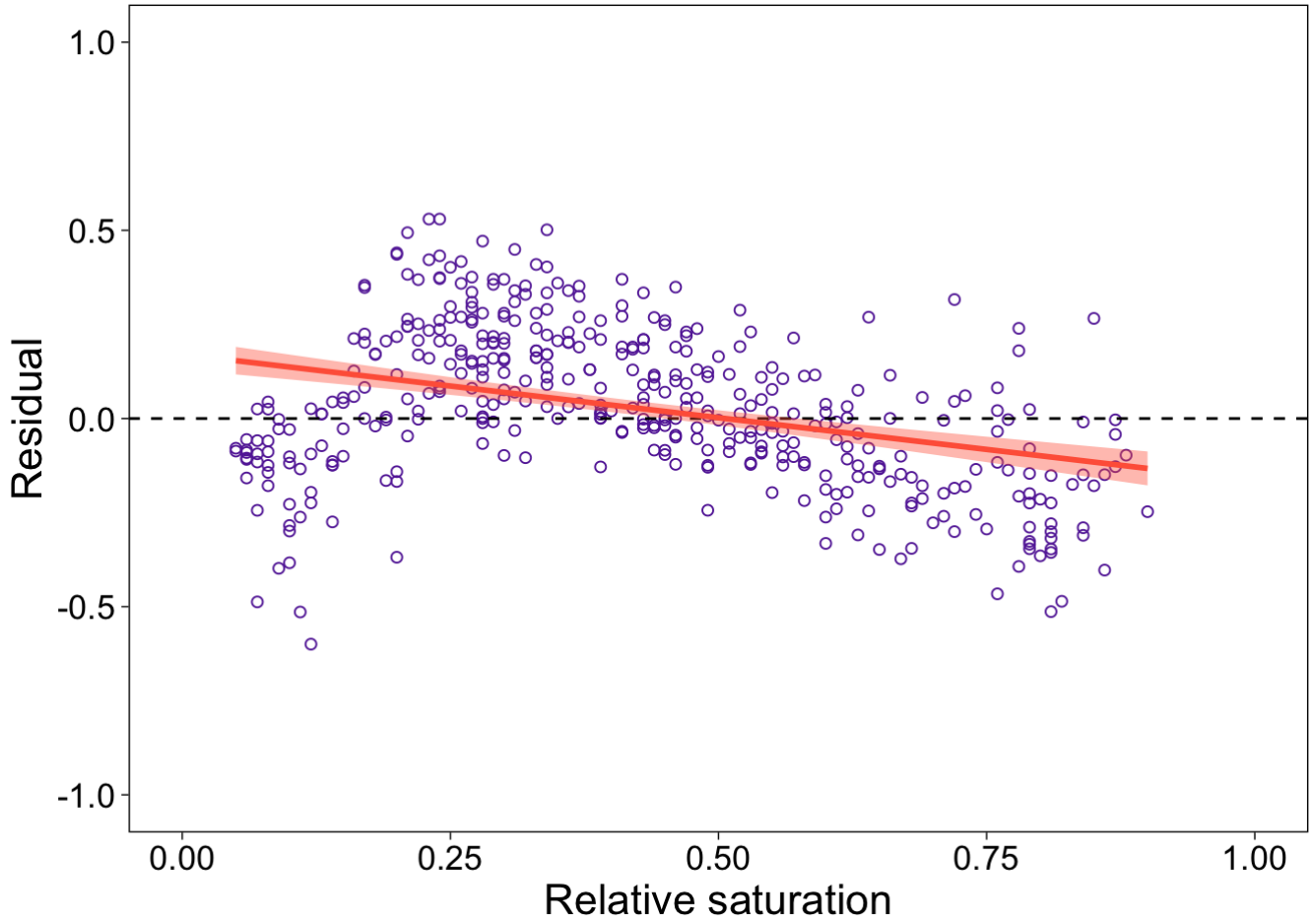
```
## [1] "R-square = 0.0836677896431547"
```

```
simple.fit = lm(res~rs, data=ptest)
summary(simple.fit)
```

```
##
## Call:
## lm(formula = res ~ rs, data = ptest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.72984 -0.11493 -0.00569  0.13457  0.44541
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.17064    0.02049   8.326 1.22e-15 ***
## rs          -0.33699    0.04338  -7.769 6.26e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.194 on 416 degrees of freedom
## Multiple R-squared:  0.1267, Adjusted R-squared:  0.1246
## F-statistic: 60.35 on 1 and 416 DF,  p-value: 6.257e-14
```

```
#residual plot
plot1<-ggplot(ptest, aes(x = rs,y = res)) +
  scale_x_continuous(name = expression(paste("Relative saturation")),limits = c(0,1)) +
  scale_y_continuous(name = 'Residual',limits = c(-1,1))+geom_point(col="#642ba6", pch=1
, cex=1.6)+geom_hline(yintercept=0, color="black", linetype="dashed",size=0.5)+
  geom_smooth(method=lm, color="tomato1", size=1, fill="tomato1", se=TRUE)

plot2<-plot1+theme_linedraw()+theme(panel.grid.major = element_blank(), panel.grid.minor
= element_blank())+theme(text = element_text(size=16))
print(plot2)
```

```
## `geom_smooth()` using formula 'y ~ x'
```


1.0
0.5
0.0
-0.5
-1.0
Residual
0.00
0.25
0.50
0.75
1.00
Relative saturation