library(ggplot2)

# read the data V4 has modified vegetation types (in site.desc variable)

hf <- read.csv ( "HF\_soil\_CO2\_v8\_rawdata\_V4.csv", header = TRUE )

hf$Tsoil.1 <- as.numeric ( as.character ( hf$Tsoil.1 ) )

hf$co2.flux[hf$co2.flux <= 0 ] <- NA

tmp <- split ( hf, paste(hf$x.SPMNAD83, hf$y.SPMNAD83, hf$collar) )

n.coll <- length(tmp)

# sort ( unlist ( lapply ( tmp, nrow) ) ) # n.obs for each collar

toosmall <- 9

toobig <- 100000

bycollar <- list()

for ( i in seq(along=tmp) ) {

# print ( paste ( "i =", i, "size =", nrow ( tmp[[i]] ) ) )

bycollar[[i]] <- list (

data = subset ( tmp[[i]], select = c ( "Tsoil.1", "co2.flux" ) ),

soil = tmp[[i]]$soil.series[1],

drainage = tmp[[i]]$Soil.drainage.class[1],

veg.type = tmp[[i]]$site.desc[1]

)

if ( nrow ( bycollar[[i]]$data ) > toosmall && nrow ( bycollar[[i]]$data ) < toobig )

bycollar[[i]]$linfit <- try ( lm ( log(co2.flux) ~ Tsoil.1,

data = bycollar[[i]]$data

),

silent = TRUE

)

else bycollar[[i]]$linfit <- "too many or too few rows"

if ( class ( bycollar[[i]]$linfit ) == "lm" ) {

bycollar[[i]]$Q10 <- exp ( 10 \* bycollar[[i]]$linfit$coef[2] )

bycollar[[i]]$R10 <- exp ( predict ( bycollar[[i]]$linfit,

newdata = data.frame ( Tsoil.1 = 10 )

)

)

}

}

rm ( tmp )

qr10 <- data.frame ( q10 = rep(NA,n.coll),

r10 = rep(NA,n.coll),

soil = factor(NA,levels=levels(hf$soil.series)),

drainage = factor(NA,levels=levels(hf$Soil.drainage.class)),

veg.type = factor(NA,levels=levels(hf$site.desc))

)

for ( i in 1:n.coll ) {

if ( class ( bycollar[[i]]$linfit ) == "lm" ) {

qr10$q10[i] <- bycollar[[i]]$Q10

qr10$r10[i] <- bycollar[[i]]$R10

qr10$soil[i] <- bycollar[[i]]$soil

qr10$drainage[i] <- bycollar[[i]]$drainage

qr10$veg.type[i] <- bycollar[[i]]$veg.type

}

}

# output data sets

outputs <- cbind(qr10$q10, qr10$r10, qr10$veg.type)

write.table(outputs, file="", sep=",")