TopdownApproach

Jinshi

May 19, 2019

# Limitation  
# 1. Savanna and grassland more than 20% area, but sample numbers are too small for Frrot, RC, and Ra/Gpp ratio  
# 2. Ra < 0 how to resolve this

# install.packages('kableExtra')  
# Load required packages  
library(cowplot)

## Loading required package: ggplot2

##   
## Attaching package: 'cowplot'

## The following object is masked from 'package:ggplot2':  
##   
## ggsave

library(data.table)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

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

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

library(ggplot2)  
theme\_set(theme\_bw())  
library(lubridate)

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:data.table':  
##   
## hour, isoweek, mday, minute, month, quarter, second, wday,  
## week, yday, year

## The following object is masked from 'package:base':  
##   
## date

library(kableExtra)

##   
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':  
##   
## group\_rows

library(cowplot)  
library(knitr)  
library("ggpubr")

## Loading required package: magrittr

##   
## Attaching package: 'ggpubr'

## The following object is masked from 'package:cowplot':  
##   
## get\_legend

library(reshape)

##   
## Attaching package: 'reshape'

## The following object is masked from 'package:lubridate':  
##   
## stamp

## The following object is masked from 'package:dplyr':  
##   
## rename

## The following object is masked from 'package:data.table':  
##   
## melt

library(ggplot2)  
# install.packages("ggmap")  
library(ggmap)

## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.

## Please cite ggmap if you use it! See citation("ggmap") for details.

##   
## Attaching package: 'ggmap'

## The following object is masked from 'package:magrittr':  
##   
## inset

## The following object is masked from 'package:cowplot':  
##   
## theme\_nothing

# install.packages("maps")  
library(maps)  
# install.packages("mapdata")  
library(mapdata)  
library(tidyr)

##   
## Attaching package: 'tidyr'

## The following objects are masked from 'package:reshape':  
##   
## expand, smiths

## The following object is masked from 'package:magrittr':  
##   
## extract

library(hexbin)  
# Source all needed functions  
  
source('functions.R')

# input data  
srdb\_v4 <- read.csv(paste0(DATA\_DIR,'/', 'srdbv4.csv'), header = T) # SRDB\_V4 data  
FlFsFr <- read.csv(paste0(DATA\_DIR,'/', 'FlFsFrSummary.csv'), header = T)  
GPP <- read.csv(paste0(DATA\_DIR,'/', 'GlobalGPP\_Sum.csv'), header = T)  
NPP <- read.csv(paste0(DATA\_DIR,'/', 'ITONPP.CSV'), header = T)  
GlobalRs <- read.csv(paste0(DATA\_DIR,'/', 'GlobalRs.csv'), header = T)  
RaGPP <- read.csv(paste0(DATA\_DIR,'/', 'RaGPP.csv'), header = T)  
IGBP <- read.table(paste0(DATA\_DIR,'/', 'IGBP.txt'), header = T, sep = ",")  
# sort(unique(IGBP$IGBP2001Pr))  
# there are 16 IGBP group  
IGBP %>% mutate(IGBP = case\_when( IGBP2001Pr == 0 ~ "WAT",  
 IGBP2001Pr == 1 ~ "ENF",  
 IGBP2001Pr == 2 ~ "EBF",  
 IGBP2001Pr == 3 ~ "DNF",  
 IGBP2001Pr == 4 ~ "DBF",  
 IGBP2001Pr == 5 ~ "MF",  
 IGBP2001Pr == 6 ~ "CSH",  
 IGBP2001Pr == 7 ~ "OSH",  
 IGBP2001Pr == 8 ~ "WSA",  
 IGBP2001Pr == 9 ~ "SAV",  
 IGBP2001Pr == 10 ~ "GRA",  
 IGBP2001Pr == 11 ~ "WET",  
 IGBP2001Pr == 12 ~ "CRO",  
 IGBP2001Pr == 13 ~ "URB",  
 IGBP2001Pr == 14 ~ "CVM",  
 IGBP2001Pr == 15 ~ "SNO",  
 IGBP2001Pr == 16 ~ "BSV",  
 TRUE ~ "UNC"  
 ) ) -> IGBP  
  
# 16 IGBP group into 10 Ecosystems  
IGBP %>% mutate (Ecosystem = case\_when( IGBP2001Pr == 0 ~ "Water",  
 IGBP2001Pr == 1 ~ "Forest",  
 IGBP2001Pr == 2 ~ "Forest",  
 IGBP2001Pr == 3 ~ "Forest",  
 IGBP2001Pr == 4 ~ "Forest",  
 IGBP2001Pr == 5 ~ "Forest",  
 IGBP2001Pr == 6 ~ "Shrubland",  
 IGBP2001Pr == 7 ~ "Shrubland",  
 IGBP2001Pr == 8 ~ "Savanna",  
 IGBP2001Pr == 9 ~ "Savanna",  
 IGBP2001Pr == 10 ~ "Grassland",  
 IGBP2001Pr == 11 ~ "Wetland",  
 IGBP2001Pr == 12 ~ "Agriculture",  
 IGBP2001Pr == 13 ~ "Urbland",  
 IGBP2001Pr == 14 ~ "Agriculture",  
 IGBP2001Pr == 15 ~ "Snow",  
 IGBP2001Pr == 16 ~ "Desert",  
 TRUE ~ "Unclassified")) -> IGBP  
# DGRsD <- read.csv(paste0(DATA\_DIR,'/', 'DGRsD.csv'), header = T)

# IGBP area by type  
IGBP %>% filter(Ecosystem != "Water" & Ecosystem != "Snow" & Ecosystem != "Unclassified" & Ecosystem != "Urbland") %>%   
 group\_by (Ecosystem) %>% summarise ( count = n() ) %>% mutate (percent = count/sum(count))

## # A tibble: 7 x 3  
## Ecosystem count percent  
## <chr> <int> <dbl>  
## 1 Agriculture 8171 0.191   
## 2 Desert 7269 0.170   
## 3 Forest 9132 0.214   
## 4 Grassland 6545 0.153   
## 5 Savanna 6628 0.155   
## 6 Shrubland 4850 0.113   
## 7 Wetland 173 0.00405

# sum(0.191,0.17,0.214,0.153,0.155,0.113,0.004)  
  
# Remove desert and shrubland because no data for Froot and Ra/GPP ratio  
IGBP %>% filter(Ecosystem != "Water" & Ecosystem != "Snow" & Ecosystem != "Unclassified" & Ecosystem != "Urbland" & Ecosystem != "Desert" & Ecosystem != "Shrubland") %>%   
 group\_by (Ecosystem) %>% summarise ( count = n() ) %>% mutate (percent = count/sum(count))

## # A tibble: 5 x 3  
## Ecosystem count percent  
## <chr> <int> <dbl>  
## 1 Agriculture 8171 0.267   
## 2 Forest 9132 0.298   
## 3 Grassland 6545 0.214   
## 4 Savanna 6628 0.216   
## 5 Wetland 173 0.00564

sum(0.267,0.298,0.213,0.216,0.006)

## [1] 1

# plot Ra Rs relationship  
# Ra\_Rs\_relationship <- function (sdata) {  
# sdata$Ra\_Norm <- ifelse (is.na(sdata$Ra\_Norm), sdata$RS\_Norm - sdata$Rh\_Norm, sdata$Ra\_Norm)  
# sdata %>% filter (!is.na(Ra\_Norm) & !is.na(RS\_Norm) & !is.na(Ecosystem\_type)  
# & Ecosystem\_type != 'Forest') -> sdata  
# ggplot(data = sdata, aes(Ra\_Norm, RS\_Norm)) + geom\_point() + facet\_wrap(~Ecosystem\_type)  
# }  
  
# Ra\_Rs\_relationship(DGRsD)

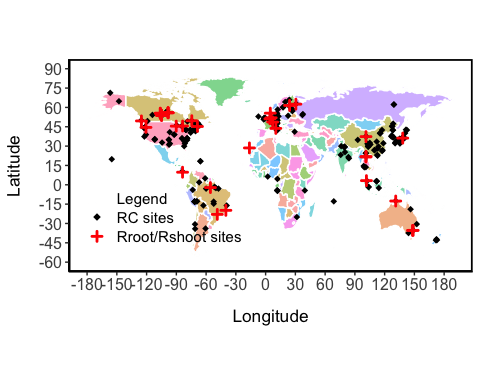
* 1. Introduction

\*\* Figure. Sites distribution – FlFsFr sites and RC/HC sites

# plot sites distribution of MGRsD  
plot\_sites(srdb\_v4, FlFsFr)

## Warning in qt(conf.interval/2 + 0.5, datac$N - 1): NaNs produced

## Warning: Removed 1 rows containing missing values (geom\_point).



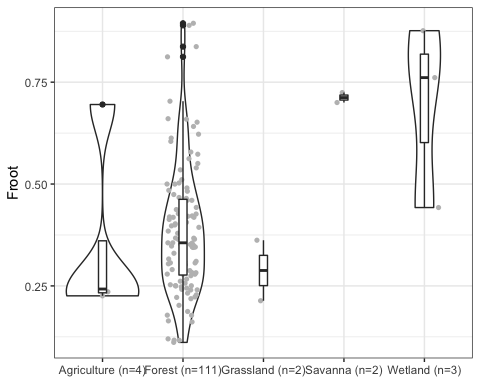
\*\* Figure, plot FfFrFw

# potential issue: only two points from grassland  
# no point from tundra   
Froot <- cal\_Froot(FlFsFr, srdb\_v4)  
  
plot\_froot <- function (sdata) {  
 var\_obs <- nrow(sdata)  
 # plot  
 Fl\_plot <- ggplot(sdata, aes(x = Ecosystem, y=Froot)) + geom\_violin() +  
 geom\_jitter(shape=16, position=position\_jitter(0.2), col = "gray") +  
 geom\_boxplot(width=.1) +  
 ylab("Froot") +  
 # stat\_summary(fun.y=median, geom="point", size=2, color="red") +  
 scale\_x\_discrete(limits = c("Agriculture", "Forest", "Grassland", "Savanna", "Wetland"),  
 labels = c("Agriculture (n=4)", "Forest (n=111)", "Grassland (n=2)", "Savanna (n=2)", "Wetland (n=3)") ) +  
 theme(axis.title.x=element\_blank())  
   
 print(median(sdata$Fr, na.rm = T) / (median(sdata$Fl, na.rm = T)+median(sdata$Fs, na.rm = T)+median(sdata$Fr, na.rm = T))\*100)  
 se <- sd(sdata$Fr, na.rm = T)/nrow(subset(sdata, !is.na(sdata$Fr)))  
 paste0("se=", se) %>% print()  
 CI <- round(qt(0.975,df=nrow(subset(sdata, !is.na(sdata$Fr)))-1)\*se, 3)  
 paste0("Fr 95% CI=", CI," ,obs=", nrow(subset(sdata, !is.na(sdata$Fr)))) %>% print()  
   
 print(median(sdata$Fs, na.rm = T) / (median(sdata$Fl, na.rm = T)+median(sdata$Fs, na.rm = T)+median(sdata$Fr, na.rm = T))\*100)  
 se <- sd(sdata$Fs, na.rm = T)/nrow(subset(sdata, !is.na(sdata$Fs)))  
 paste0("se=", se) %>% print()  
 CI <- round(qt(0.975,df=nrow(subset(sdata, !is.na(sdata$Fs)))-1)\*se, 3)  
 paste0("Fs 95% CI=", CI, " ,obs=", nrow(subset(sdata, !is.na(sdata$Fs)))) %>% print()  
   
 print(median(sdata$Fl, na.rm = T) / (median(sdata$Fl, na.rm = T)+median(sdata$Fs, na.rm = T)+median(sdata$Fr, na.rm = T))\*100)  
 se <- sd(sdata$Fl, na.rm = T)/nrow(subset(sdata, !is.na(sdata$Fl)))  
 paste0("se=", se) %>% print()  
 CI <- round(qt(0.975,df=nrow(subset(sdata, !is.na(sdata$Fl)))-1)\*se,3)   
 paste0("Fl 95% CI=", CI ," ,obs=", nrow(subset(sdata, !is.na(sdata$Fl)))) %>% print()  
   
 print(Fl\_plot)  
}  
  
# check whether Froot + Fshoot = 1  
# Froot %>% transmute (check = Froot + Fshoot) %>% max  
# Froot %>% group\_by(Ecosystem) %>% summarise(count = n())   
# mean(Froot$Froot)  
plot\_froot (Froot)

## numeric(0)  
## [1] "se=0.00141521617936181"  
## [1] "Fr 95% CI=0.003 ,obs=122"  
## numeric(0)  
## [1] "se=0.00141521617936181"  
## [1] "Fs 95% CI=0.003 ,obs=122"  
## numeric(0)  
## [1] "se=NA"

## Warning in qt(0.975, df = nrow(subset(sdata, !is.na(sdata$Fl))) - 1): NaNs  
## produced

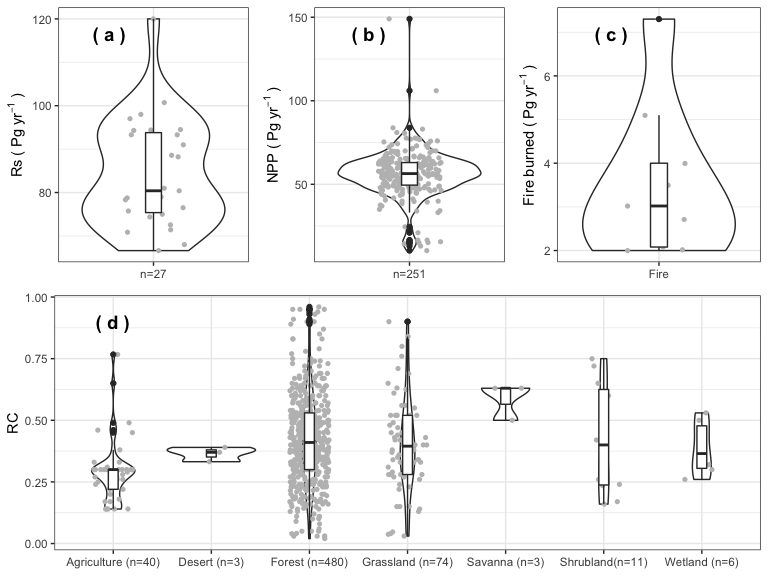
## [1] "Fl 95% CI=NaN ,obs=0"



\*\* Figure, plot global Rs

plot\_Rs (GlobalRs, srdb\_v4, NPP)

## # A tibble: 7 x 2  
## Ecosystem\_type n  
## <fct> <int>  
## 1 Agriculture 40  
## 2 Desert 3  
## 3 Forest 480  
## 4 Grassland 74  
## 5 Savanna 3  
## 6 Shrubland 11  
## 7 Wetland 6  
## [1] 0.4121125  
## [1] "RC se=0.00740339660633126"  
## [1] "RC 95% CI=0.015"  
## [1] "RC obs=617"  
## [1] 56.4  
## [1] "NPP se=0.902534118625694"  
## [1] "NPP 95% CI=1.778"

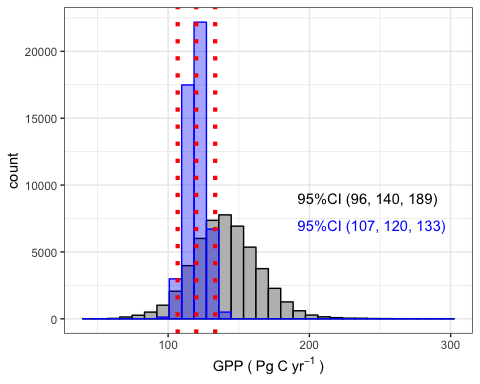


# Bottom-up approache to stimate GPP  
# Rroot = Rs \* RC (Rroot to Rs ratio, from srdb)  
# Froot = Rroot / Ra (root autotrophic respiration to total autotrophic respiration fraction/ratio, calculation please see plot\_froot function)  
# Fshoot = 1-Froot  
# FsFr = (1-Froot)/Froot (FsFr = Rshoot/Rroot ratio)  
# Ra = Rroot + Rshoot  
  
var\_RC <- srdb\_v4 %>% select(RC\_annual, Ecosystem\_type) %>% filter(RC\_annual > 0.01 & RC\_annual < 0.99 & !is.na(RC\_annual))   
var\_agr <- var\_RC %>% filter (Ecosystem\_type=="Agriculture")   
var\_des <- var\_RC %>% filter (Ecosystem\_type=="Desert")   
var\_for <- var\_RC %>% filter (Ecosystem\_type=="Forest")   
var\_gra <- var\_RC %>% filter (Ecosystem\_type=="Grassland")   
var\_sav <- var\_RC %>% filter (Ecosystem\_type=="Savanna")  
var\_shr <- var\_RC %>% filter (Ecosystem\_type=="Shrubland")   
var\_wet <- var\_RC %>% filter (Ecosystem\_type=="Wetland")   
  
var\_Rs <- GlobalRs %>% filter(!is.na(Rs) & is.na(SD) )  
  
n\_samp <- 50000  
n\_rep <- 50000  
  
gpp\_results <- tibble(  
 # Rs = replicate( n\_rep, sample(var\_Rs$Rs, n\_samp, replace = TRUE) %>% mean() ),  
 # Rs1 = rnorm(n = n\_samp, mean = 68, sd = 4),  
 Rs2 = rnorm(n = n\_samp, mean = 80.4, sd = 16.9),  
 Rs3 = rnorm(n = n\_samp, mean = 98, sd = 12),  
 Rs4 = rnorm(n = n\_samp, mean = 79, sd = 7.65),  
 Rs5 = rnorm(n = n\_samp, mean = 94.4, sd = 4.59),  
 Rs6 = rnorm(n = n\_samp, mean = 91, sd = 2.05),  
 Rs7 = rnorm(n = n\_samp, mean = 94.3, sd = 9.13),  
 Rs8 = rnorm(n = n\_samp, mean = 78.34, sd = 2.25),  
 Rs9 = rnorm(n = n\_samp, mean = 72.55, sd = 7.13),  
 Rs10 = sample(var\_Rs$Rs, n\_samp, replace = TRUE),  
 Rs = (Rs2 + Rs3 + Rs4 + Rs5 + Rs6 + Rs7 + Rs8 + Rs9 + Rs10)/9,  
   
 # Rc = replicate( n\_rep, var\_RC$RC\_annual, n\_samp, replace = TRUE) %>% mean() ),  
 Rc\_agr = sample(var\_agr$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_des = sample(var\_des$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_for = sample(var\_for$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_gra = sample(var\_gra$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_sav = sample(var\_sav$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_shr = sample(var\_shr$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc\_wet = sample(var\_wet$RC\_annual, n\_samp, replace = TRUE) ,  
 Rc = (Rc\_agr\*0.191 + Rc\_des\*0.17 + Rc\_for\*0.214 + Rc\_gra\*0.153 + Rc\_sav\*0.155 + Rc\_shr\*0.113 + Rc\_wet\*0.004),  
 # Froot = sample(Froot$Froot, n\_samp, replace = TRUE) ,  
 Froot\_agr = sample(Froot[Froot$Ecosystem == "Agriculture",]$Froot, n\_samp, replace = TRUE),  
 Froot\_for = sample(Froot[Froot$Ecosystem == "Forest",]$Froot, n\_samp, replace = TRUE),  
 Froot\_gra = sample(Froot[Froot$Ecosystem == "Grassland",]$Froot, n\_samp, replace = TRUE),  
 Froot\_sav = sample(Froot[Froot$Ecosystem == "Savanna",]$Froot, n\_samp, replace = TRUE),  
 Froot\_wet = sample(Froot[Froot$Ecosystem == "Wetland",]$Froot, n\_samp, replace = TRUE),  
 Froot = Froot\_agr\*0.267 + Froot\_for\*0.298 + Froot\_gra\*0.213 + Froot\_sav\*0.216 + Froot\_wet\*0.006,  
 # FsFr = 1 / (1 - Froot),  
 FsFr = (1 - Froot) / Froot,  
 NPP = sample(NPP$NPP, n\_samp, replace = TRUE) ,  
   
 gpp1 = rnorm(n = n\_samp, mean = 124.6, sd = 2.7),  
 gpp2 = rnorm(n = n\_samp, mean = 135.7, sd = 21.712),  
 gpp3 = rnorm(n = n\_samp, mean = 109.29, sd = 27.33),  
 gpp4 = rnorm(n = n\_samp, mean = 118, sd = 26),  
 gpp5 = rnorm(n = n\_samp, mean = 119, sd = 6),  
 gpp6 = rnorm(n = n\_samp, mean = 110.5, sd = 21.3),  
 gpp7 = rnorm(n = n\_samp, mean = 117, sd = 13),  
 gpp8 = sample(GPP[is.na(GPP$SD),]$GPP, n\_samp, replace = TRUE),  
 GPP\_raw = (gpp1 + gpp2 + gpp3 + gpp4 + gpp5 + gpp6 + gpp7 + gpp8)/8,  
 Rroot = Rs \* Rc,  
 Rshoot = Rroot \* FsFr,  
 GPP = NPP + Rroot + Rshoot  
)

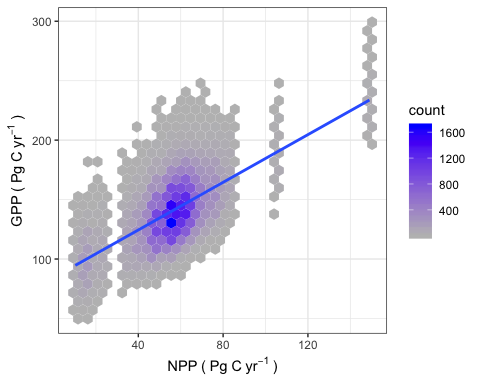
# check  
gpp\_results %>% mutate(check = (1-Froot)/Froot, FsFr)

## # A tibble: 50,000 x 39  
## Rs2 Rs3 Rs4 Rs5 Rs6 Rs7 Rs8 Rs9 Rs10 Rs Rc\_agr  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 64.2 88.6 99.2 98.3 91.7 88.6 78.6 68.3 76.2 83.7 0.45  
## 2 67.8 107. 64.5 94.5 91.3 87.1 75.4 80.1 66.6 81.7 0.27  
## 3 86.6 95.9 83.5 97.1 91.3 98.0 77.4 64.2 78.8 85.9 0.14  
## 4 48.5 101. 74.9 100. 89.6 102. 75.5 61.3 68 80.2 0.24  
## 5 84.2 93.9 90.3 100.0 90.5 97.0 80.5 74.5 74.4 87.2 0.17  
## 6 75.7 102. 72.7 88.0 87.7 95.4 74.9 71.8 88.6 84.1 0.17  
## 7 80.6 97.3 86.3 95.7 91.6 86.0 77.1 72.6 76.2 84.8 0.14  
## 8 67.2 121. 92.7 94.1 87.8 95.5 78.0 65.4 94.5 88.5 0.33  
## 9 55.3 113. 71.6 93.0 93.9 98.6 76.2 71.7 81.0 83.8 0.14  
## 10 68.1 92.5 65.2 95.5 91.8 91.2 76.1 73.7 74.4 80.9 0.14  
## # … with 49,990 more rows, and 28 more variables: Rc\_des <dbl>,  
## # Rc\_for <dbl>, Rc\_gra <dbl>, Rc\_sav <dbl>, Rc\_shr <dbl>, Rc\_wet <dbl>,  
## # Rc <dbl>, Froot\_agr <dbl>, Froot\_for <dbl>, Froot\_gra <dbl>,  
## # Froot\_sav <dbl>, Froot\_wet <dbl>, Froot <dbl>, FsFr <dbl>, NPP <dbl>,  
## # gpp1 <dbl>, gpp2 <dbl>, gpp3 <dbl>, gpp4 <dbl>, gpp5 <dbl>,  
## # gpp6 <dbl>, gpp7 <dbl>, gpp8 <dbl>, GPP\_raw <dbl>, Rroot <dbl>,  
## # Rshoot <dbl>, GPP <dbl>, check <dbl>

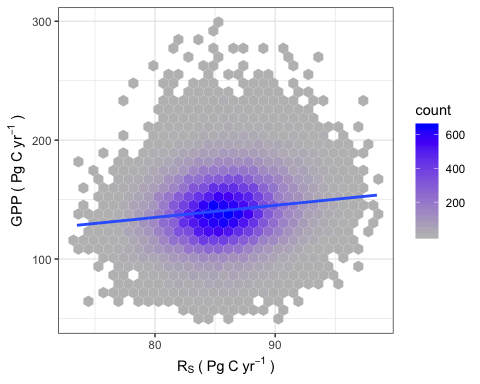
# hist(gpp\_results$GPP) + abline (v = quantile (gpp\_results$GPP, c(0.025, 0.5, 0.975)), col = "red", lty = "dashed")  
# mean(gpp\_results$GPP)  
# max(gpp\_results$GPP)  
  
var\_CI <- quantile(gpp\_results$GPP, c(0.025, 0.5, 0.975)) %>% round(0)  
var\_CI\_raw <- quantile(gpp\_results$GPP\_raw, c(0.025, 0.5, 0.975)) %>% round(0)  
ggplot(gpp\_results) + aes(GPP) + geom\_histogram(col = "black", fill = "gray", bins = 30) +   
 geom\_histogram(aes(GPP\_raw), col = "blue", fill = "blue", bins = 30, alpha = 0.35) +  
 geom\_vline(xintercept = c(quantile (gpp\_results$GPP\_raw, c(0.025, 0.5, 0.975))),col = "red", linetype = "dotted", size = 1.5) +  
 annotate("text", x = quantile(gpp\_results$GPP, 0.98), y = 9000, label = paste0("95%CI (", var\_CI[1],", ", var\_CI[2], ", ", var\_CI[3], ")"), hjust = 0) +  
 annotate("text", x = quantile(gpp\_results$GPP, 0.98), y = 7000, label = paste0("95%CI (", var\_CI\_raw[1],", ", var\_CI\_raw[2], ", ", var\_CI\_raw[3], ")")  
 , hjust = 0, col = "blue") +  
 xlab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



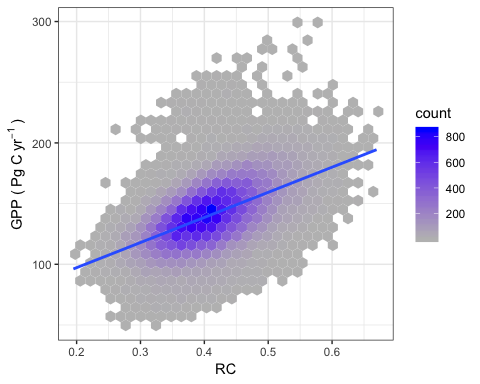
# GPP ~ NPP  
ggplot(gpp\_results) +  
 aes(x = NPP, y = GPP) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +   
 xlab(expression(NPP~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



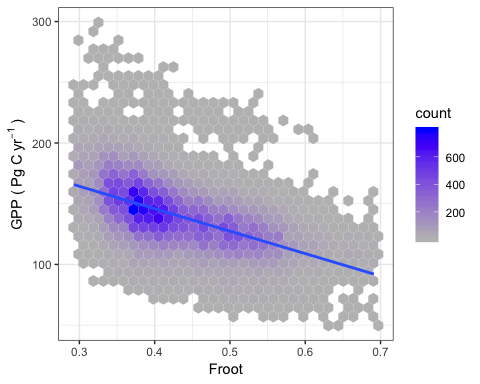
# GPP ~ Rs  
ggplot(gpp\_results) +  
 aes(x = Rs, y = GPP) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(R[S]~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



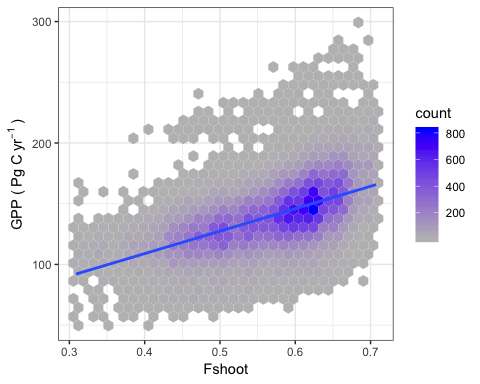
# GPP ~ Rc  
ggplot(gpp\_results) +  
 aes(x = Rc, y = GPP) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(RC)) +  
 ylab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



# GPP ~ Froot  
ggplot(gpp\_results) +  
 aes(x = Froot, y = GPP) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(Froot)) +  
 ylab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



# GPP ~ Fshoot  
ggplot(gpp\_results) +  
 aes(x = 1-Froot, y = GPP) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(Fshoot)) +  
 ylab(expression(GPP~"("~Pg~C~yr^{-1}~")"))



# Quantify the properbility GPP following (107, 133)  
# about 29% chance GPP will between (107-133)  
y <- mean(gpp\_results$GPP)  
s <- sd(gpp\_results$GPP)  
p\_gpp\_less\_than\_133 <- pnorm ((133 - y)/(s))  
p\_gpp\_less\_than\_107 <- pnorm ((107 - y)/(s))  
p\_gpp <- p\_gpp\_less\_than\_133 -p\_gpp\_less\_than\_107  
p\_gpp

## [1] 0.2944065

# pnorm (-0.329)

\*\* Figure. GPP plot

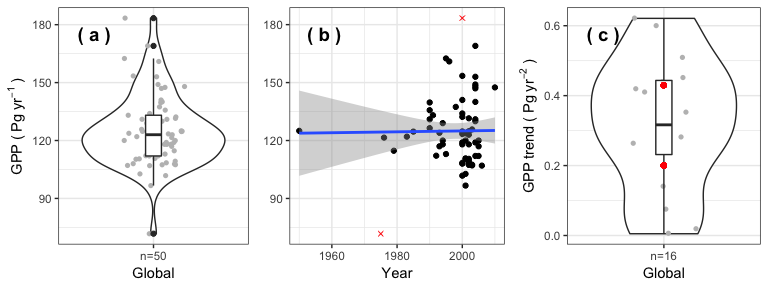
# The red dots in panel c are calculated trend from panel b (increase rate = 0.2 if outliers removed, increase rate = 0.42 if outlers kept)  
plot\_GPP(GPP)

## [1] "median=123"  
## [1] "se=2.29621351780485"  
## [1] "95% CI=4.587"  
## [1] "number of trend=16"  
##   
## Call:  
## lm(formula = GPP ~ Year, data = subGPP)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.327 -12.859 -2.018 7.810 43.935   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 77.9609 452.1737 0.172 0.864  
## Year 0.0235 0.2263 0.104 0.918  
##   
## Residual standard error: 16.03 on 61 degrees of freedom  
## Multiple R-squared: 0.0001768, Adjusted R-squared: -0.01621   
## F-statistic: 0.01079 on 1 and 61 DF, p-value: 0.9176  
##   
##   
## Call:  
## lm(formula = GPP ~ Year, data = GPP)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.315 -14.685 -2.175 9.840 57.759   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -401.2559 496.2209 -0.809 0.422  
## Year 0.2634 0.2484 1.061 0.293  
##   
## Residual standard error: 18.49 on 63 degrees of freedom  
## Multiple R-squared: 0.01754, Adjusted R-squared: 0.001946   
## F-statistic: 1.125 on 1 and 63 DF, p-value: 0.2929

## Warning: Removed 49 rows containing non-finite values (stat\_ydensity).

## Warning: Removed 49 rows containing non-finite values (stat\_boxplot).

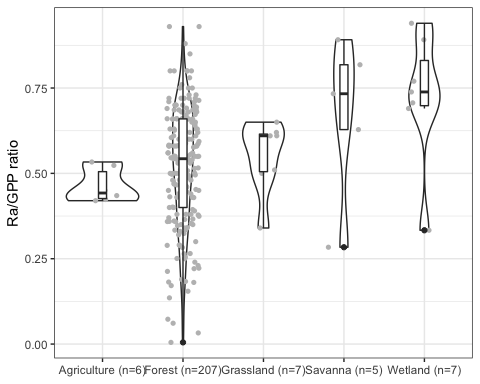
## Warning: Removed 49 rows containing missing values (geom\_point).



\*\* Figure plot Ra to GPP ratio

# only one point from tundra  
var\_RaGpp <- plot\_RaGPP (RaGPP, srdb\_v4)

## [1] 0.5303451  
## [1] "Ra/GPP se=0.0117196199746696"  
## [1] "Ra/GPP 95% CI=0.023"  
## [1] "Ra/GPP obs=232"



var\_RaGpp %>% group\_by (Ecosystem) %>% summarise ( count = n() ) %>% mutate (percent = count/sum(count)) %>% arrange (desc(Ecosystem))

## # A tibble: 5 x 3  
## Ecosystem count percent  
## <fct> <int> <dbl>  
## 1 Savanna 5 0.0216  
## 2 Agriculture 6 0.0259  
## 3 Wetland 7 0.0302  
## 4 Grassland 7 0.0302  
## 5 Forest 207 0.892

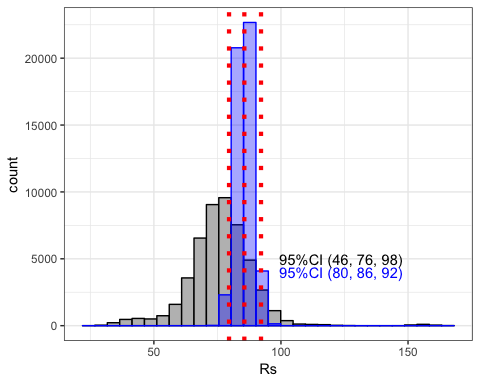
# topdowm approach to eatimate Rs  
# Ra = GPP \* Ra\_GPP\_ratio (calculated based on data from a summary paper and from srdb, see plot\_RaGPP function)  
# Ra = GPP - NPP (two approaches yield a very similar results, we thus used the average of two estimates)  
  
var\_Fire <- c(2, 3.5, 7.3, 4, 5.1, 2.02, 2.71, 3.02, 2.08)  
n\_samp <- 50000  
# n\_rep <- 3  
  
rs\_results <- tibble(  
 # GPP = replicate(n\_rep, sample(GPP$GPP, n\_samp, replace = TRUE) %>% mean() ),  
 # GPP = sample(GPP$GPP, n\_samp, replace = TRUE),  
 gpp1 = rnorm(n = n\_samp, mean = 124.6, sd = 2.7),  
 gpp2 = rnorm(n = n\_samp, mean = 135.7, sd = 21.712),  
 gpp3 = rnorm(n = n\_samp, mean = 109.29, sd = 27.33),  
 gpp4 = rnorm(n = n\_samp, mean = 118, sd = 26),  
 gpp5 = rnorm(n = n\_samp, mean = 119, sd = 6),  
 gpp6 = rnorm(n = n\_samp, mean = 110.5, sd = 21.3),  
 gpp7 = rnorm(n = n\_samp, mean = 117, sd = 13),  
 gpp8 = sample(GPP[is.na(GPP$SD),]$GPP, n\_samp, replace = TRUE),  
 GPP = (gpp1 + gpp2 + gpp3 + gpp4 + gpp5 + gpp6 + gpp7 + gpp8)/8,  
   
 NPP = sample(NPP$NPP, n\_samp, replace = TRUE),  
 HerbComsum = rnorm(n = n\_rep, mean = 2.2, sd = 0.2),  
 Fire = sample(var\_Fire, n\_samp, replace = TRUE),  
 sink = rnorm(n = n\_rep, mean = 2.10, sd = 0.28),  
   
 # RaGpp = sample(var\_RaGpp$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp\_agr = sample(var\_RaGpp[var\_RaGpp$Ecosystem == "Agriculture", ]$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp\_for = sample(var\_RaGpp[var\_RaGpp$Ecosystem == "Forest", ]$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp\_gra = sample(var\_RaGpp[var\_RaGpp$Ecosystem == "Grassland", ]$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp\_sav = sample(var\_RaGpp[var\_RaGpp$Ecosystem == "Savanna", ]$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp\_wet = sample(var\_RaGpp[var\_RaGpp$Ecosystem == "Wetland", ]$RaGPP\_ratio, n\_samp, replace = TRUE),  
 RaGpp = RaGpp\_agr\*0.267 + RaGpp\_for\*0.298 + RaGpp\_gra\*0.213 + RaGpp\_sav\*0.216 + RaGpp\_wet\*0.006,  
   
 # Froot = sample(Froot$Froot, n\_samp, replace = TRUE) ,  
 Froot\_agr = sample(Froot[Froot$Ecosystem == "Agriculture",]$Froot, n\_samp, replace = TRUE),  
 Froot\_for = sample(Froot[Froot$Ecosystem == "Forest",]$Froot, n\_samp, replace = TRUE),  
 Froot\_gra = sample(Froot[Froot$Ecosystem == "Grassland",]$Froot, n\_samp, replace = TRUE),  
 Froot\_sav = sample(Froot[Froot$Ecosystem == "Savanna",]$Froot, n\_samp, replace = TRUE),  
 Froot\_wet = sample(Froot[Froot$Ecosystem == "Wetland",]$Froot, n\_samp, replace = TRUE),  
 Froot = Froot\_agr\*0.267 + Froot\_for\*0.298 + Froot\_gra\*0.213 + Froot\_sav\*0.216 + Froot\_wet\*0.006,  
 Fshoot = 1 - Froot,  
   
 # Rs\_raw = sample(na.omit(GlobalRs$Rs), n\_samp, replace = TRUE) ,  
 Rs2 = rnorm(n = n\_samp, mean = 80.4, sd = 16.9),  
 Rs3 = rnorm(n = n\_samp, mean = 98, sd = 12),  
 Rs4 = rnorm(n = n\_samp, mean = 79, sd = 7.65),  
 Rs5 = rnorm(n = n\_samp, mean = 94.4, sd = 4.59),  
 Rs6 = rnorm(n = n\_samp, mean = 91, sd = 2.05),  
 Rs7 = rnorm(n = n\_samp, mean = 94.3, sd = 9.13),  
 Rs8 = rnorm(n = n\_samp, mean = 78.34, sd = 2.25),  
 Rs9 = rnorm(n = n\_samp, mean = 72.55, sd = 7.13),  
 Rs10 = sample(var\_Rs$Rs, n\_samp, replace = TRUE),  
 Rs\_raw = (Rs2 + Rs3 + Rs4 + Rs5 + Rs6 + Rs7 + Rs8 + Rs9 + Rs10)/9,  
   
 Ra1 = GPP \* RaGpp,  
 Ra2 = ifelse(GPP<NPP, 0, (GPP - NPP)),  
 Ra = (Ra1+Ra2)/2,  
   
 Rroot = Ra \* Froot,  
 Rshoot = Ra \* Fshoot,  
 Rs = NPP - HerbComsum - Fire - sink + Rroot   
)

rs\_results$Ra2 %>% min(na.rm = T)

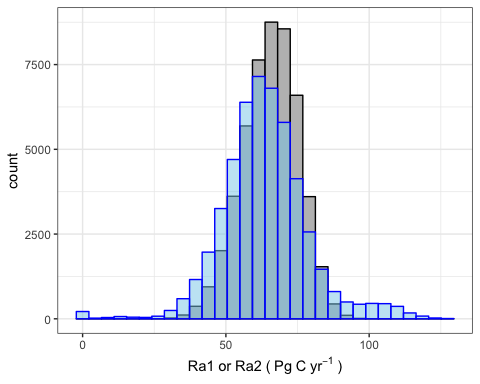
## [1] 0

### Topdow approach results diagnose

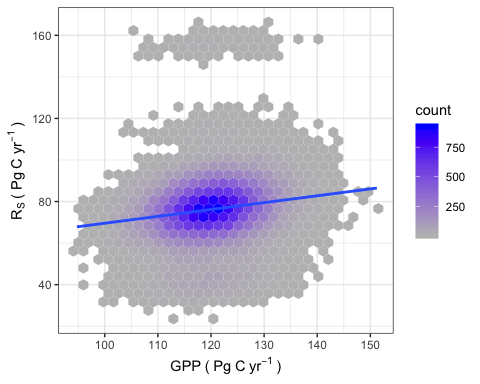
# mean(rs\_results$Rs)  
# max(rs\_results$Rs)  
  
var\_CI <- quantile(rs\_results$Rs, c(0.025, 0.5, 0.975)) %>% round(0)  
var\_CI\_raw <- quantile(rs\_results$Rs\_raw, c(0.025, 0.5, 0.975)) %>% round(0)  
ggplot(rs\_results) + aes(Rs) + geom\_histogram(col = "black", fill = "gray", bins = 30) +   
 geom\_histogram(aes(Rs\_raw), col = "blue", fill = "blue", bins = 30, alpha = 0.35) +  
 geom\_vline(xintercept = c(quantile (rs\_results$Rs\_raw, c(0.025, 0.5, 0.975))),col = "red", linetype = "dotted", size = 1.5) +  
 annotate("text", x = quantile(rs\_results$Rs, 0.98), y = 5000, label = paste0("95%CI (", var\_CI[1],", ", var\_CI[2], ", ", var\_CI[3], ")"), hjust = 0) +  
 annotate("text", x = quantile(rs\_results$Rs, 0.98), y = 4000, label = paste0("95%CI (", var\_CI\_raw[1],", ", var\_CI\_raw[2], ", ", var\_CI\_raw[3], ")")  
 , hjust = 0, col = "blue") +  
 ylab(expression(count))



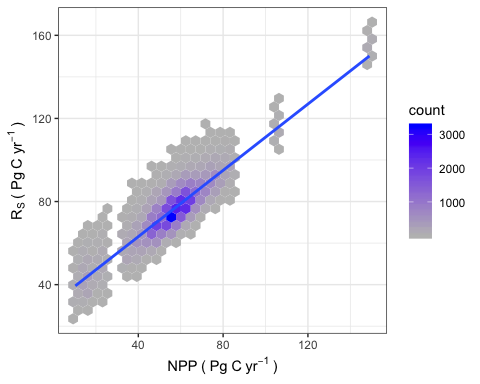
# Rs raw data plot  
# var\_CI <- quantile(rs\_results$Rs\_raw, c(0.025, 0.5, 0.975)) %>% round(0)  
# plot\_rs\_raw <- ggplot(rs\_results) + aes(Rs\_raw) + geom\_histogram(col = "black", fill = "gray", bins = 30) +   
# geom\_vline(xintercept = c(quantile (rs\_results$Rs, c(0.025, 0.5, 0.975))),col = "red", linetype = "dotted", size = 1.5) +  
# annotate("text", x = quantile(rs\_results$Rs, 0.98), y = 5000, label = paste0("95%CI (", var\_CI[1],", ", var\_CI[2], ", ", var\_CI[3], ")"), hjust = 0)  
  
# Ra1 ~ Ra2  
ggplot(rs\_results) +  
 aes(x = Ra1) + geom\_histogram(col = "black", fill = "gray", bins = 30) +  
 geom\_histogram(aes(Ra2), col = "blue", fill = "skyblue", bins = 30, alpha = 0.5) +  
 xlab(expression(Ra1~or~Ra2~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(count))



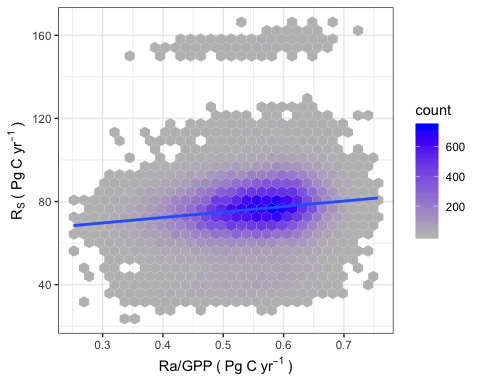
# Rs ~ GPP  
ggplot(rs\_results) +  
 aes(x = GPP, y = Rs) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(GPP~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(R[S]~"("~Pg~C~yr^{-1}~")"))



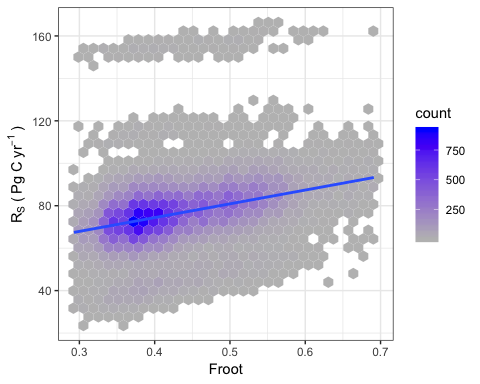
# Rs ~ NPP  
ggplot(rs\_results) +  
 aes(x = NPP, y = Rs) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(NPP~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(R[S]~"("~Pg~C~yr^{-1}~")"))



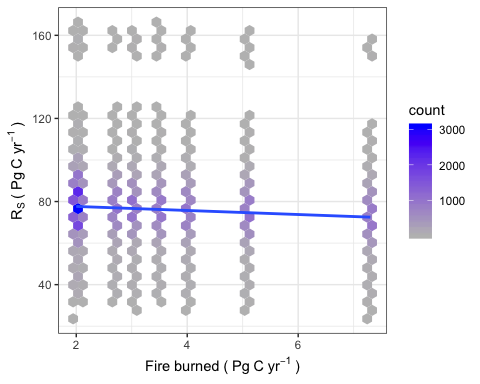
# Rs ~ RaGpp  
ggplot(rs\_results) +  
 aes(x = RaGpp, y = Rs) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression("Ra/GPP"~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(R[S]~"("~Pg~C~yr^{-1}~")"))



# Rs ~ Froot  
ggplot(rs\_results) +  
 aes(x = Froot, y = Rs) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(Froot)) +  
 ylab(expression(R[S]~"("~Pg~C~yr^{-1}~")"))



# Rs ~ Fire  
ggplot(rs\_results) +  
 aes(x = Fire, y = Rs) +  
 geom\_hex() +  
 scale\_fill\_gradient(low = "gray", high = "blue") +  
 geom\_smooth(method = "lm") +  
 xlab(expression(Fire~burned~"("~Pg~C~yr^{-1}~")")) +  
 ylab(expression(R[S]~"("~Pg~C~yr^{-1}~")"))



# about 27% chance Rs will between (80-92)  
y <- mean(rs\_results$Rs)  
s <- sd(rs\_results$Rs)  
p\_Rs\_higher\_than\_80 <- pnorm ((y - 80)/(s))  
p\_Rs\_higher\_than\_92 <- pnorm ((y - 92)/(s))  
p\_Rs <- p\_Rs\_higher\_than\_80 -p\_Rs\_higher\_than\_92  
p\_Rs

## [1] 0.2741192