# Purpose: Air Microbiome project- Script to collocate OPC data

#Authors: Enzo Acerbi, Caroline Chenard

#Created: 27 Nov 2014

#Last updated: 27 Nov 2014

if (!file.exists("data")) dir.create("data")

if (!file.exists("src")) dir.create("src")

if (!file.exists("results")) dir.create("results")

if (!file.exists("doc")) dir.create("doc")

library(ggplot2)

library(stringr)

# open the data file (remove first 10 rows)

OPC\_down <- read.csv("data/collocate\_down\_sec.csv", header=T, skip=10)

OPC\_up <- read.csv("data/collocate\_up\_sec.csv", header=T, skip=10)

str (OPC\_up)

#Convert data.and.time in a format that ggplot2 can understand

tmp = OPC\_down$Date.and.Time

tmp = levels(tmp)

tmp = strptime(tmp, "%m/%d/%Y %H:%M:%S")

OPC\_down$Date.and.Time = tmp

tmp = OPC\_up$Date.and.Time

tmp = levels(tmp)

tmp = strptime(tmp, "%m/%d/%Y %H:%M:%S")

OPC\_up$Date.and.Time = tmp

# Plot the particle number based on size into 6 different plots-- check for any outliers

ch1 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch1.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch1.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("0.3" \* mu ~ "m") + xlab("") +

theme(legend.position="top") +

scale\_colour\_brewer(palette="Set1")

#ch1

ch2 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch2.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch2.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("0.5" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#ch2

ch3 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch3.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch3.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("1.0" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#ch3

ch4 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch4.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch4.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("2.5" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#ch4

ch5 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch5.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch5.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("5.0" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#ch5

ch6 <- ggplot() + geom\_point(data = OPC\_down, aes(x=Date.and.Time, y =Ch6.Diff...., colour="down")) + geom\_point(data = OPC\_up, aes(x=Date.and.Time, y =Ch6.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("10.0" \* mu ~ "m") + xlab("Time") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#ch6

#arrange plot into 2 rows and 3 columns

library(gridExtra)

grid.arrange(ch1, ch2, ch3, ch4,ch5, ch6, ncol=2, nrow = 3)

#give info on each column.

summary (OPC\_down)

#check for outliers on graph, removed them using the following command

OPC\_down\_sub <- subset(OPC\_down, OPC\_down$Ch4.Diff....<1000 & OPC\_down$Ch5.Diff....<200 & OPC\_down$Ch6.Diff....<100)

summary (OPC\_down\_sub)

# Repeat on the OPC\_up data.

#Given that the OPC didn't start in the same time, we need to remove the first 1125 reads.

#I am not sure how to remove it based on their time, So I will use the Record. colunm

OPC\_up\_sub <- subset(OPC\_up, OPC\_up$Ch4.Diff....<1000 & OPC\_up$Ch5.Diff....<200 & OPC\_up$Ch6.Diff....<100 & OPC\_up$Record..>1125)

summary (OPC\_up\_sub)

#check that the data still look the same

sub\_ch1 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch1.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch1.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("0.3" \* mu ~ "m") + xlab("") +

theme(legend.position="top") +

scale\_colour\_brewer(palette="Set1")

sub\_ch1

sub\_ch2 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch2.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch2.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("0.5" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#sub\_ch2

sub\_ch3 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch3.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch3.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("1.0" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#sub\_ch3

sub\_ch4 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch4.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch4.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("2.5" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#sub\_ch4

sub\_ch5 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch5.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch5.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("5.0" \* mu ~ "m") + xlab("") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#sub\_ch5

sub\_ch6 <- ggplot() + geom\_point(data = OPC\_down\_sub, aes(x=Date.and.Time, y =Ch6.Diff...., colour="down")) + geom\_point(data = OPC\_up\_sub, aes(x=Date.and.Time, y =Ch6.Diff...., colour="up"))+

theme\_bw(base\_size = 14) +

ylab ("10.0" \* mu ~ "m") + xlab("Time") +

theme(legend.position="none" ) +

scale\_colour\_brewer(palette="Set1")

#sub\_ch6

#arrange plot into 2 rows and 3 columns

library(gridExtra)

grid.arrange(sub\_ch1, sub\_ch2,sub\_ch3,sub\_ch4, sub\_ch5, sub\_ch6, ncol=2, nrow = 3)

# If everything looks good. Sum for each particle size counts.

# sum for 0.3um particle counts

sum\_down1 <- sum(OPC\_down\_sub$Ch1.Diff....)

sum\_up1 <-sum(OPC\_up\_sub$Ch1.Diff....)

norm1 <- sum\_down1/ sum\_up1

norm1

# sum for 0.5um particle counts

sum\_down2 <- sum(OPC\_down\_sub$Ch2.Diff....)

sum\_up2 <-sum(OPC\_up\_sub$Ch2.Diff....)

norm2 <- sum\_down2/ sum\_up2

norm2

# sum for 1.0 um particle counts

sum\_down3 <- sum(OPC\_down\_sub$Ch3.Diff....)

sum\_up3 <-sum(OPC\_up\_sub$Ch3.Diff....)

norm3 <- sum\_down3/sum\_up3

norm3

# sum for 2.5 um particle counts

sum\_down4 <- sum(OPC\_down\_sub$Ch4.Diff....)

sum\_up4 <-sum(OPC\_up\_sub$Ch4.Diff....)

norm4 <- sum\_down4/ sum\_up4

norm4

# sum for 5.0 um particle counts

sum\_down5 <- sum(OPC\_down\_sub$Ch5.Diff....)

sum\_up5 <-sum(OPC\_up\_sub$Ch5.Diff....)

norm5 <- sum\_down5/ sum\_up5

norm5

# sum for 10.0 um particle counts

sum\_down6 <- sum(OPC\_down\_sub$Ch6.Diff....)

sum\_up6 <-sum(OPC\_up\_sub$Ch6.Diff....)

norm6 <- sum\_down6/ sum\_up6

norm6

allnorm <- c(norm1, norm2, norm3, norm4, norm5, norm6)

allnorm