

# Analysis notes

July 30, 2014

V0.0.2 PN

## 1 Data manipulation

Data were crudely transformed to input, using Mike's suggestions for fishery deveopment year.

```
require(dplyr)

## Loading required package: 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

year.table <- tbl_df(read.csv("~/Work/Dropbox/First year of assessment/V4_Final_dataset.csv",
  na.strings = c("", "NA", "#N/A"), stringsAsFactors = F))

comm.landings <- tbl_df(read.csv("~/Work/Dropbox/First year of assessment/US_comm_landings.csv",
  na.strings = c("", "NA", "#N/A"), stringsAsFactors = F))

rec.landings <- tbl_df(read.csv("~/Work/Dropbox/First year of assessment/US_rec_landings.csv",
  na.strings = c("", "NA", "#N/A"), stringsAsFactors = F))

price <- tbl_df(read.csv("~/Work/Dropbox/First year of assessment/US_price.csv",
  na.strings = c("", "NA", "#N/A"), stringsAsFactors = F))
```

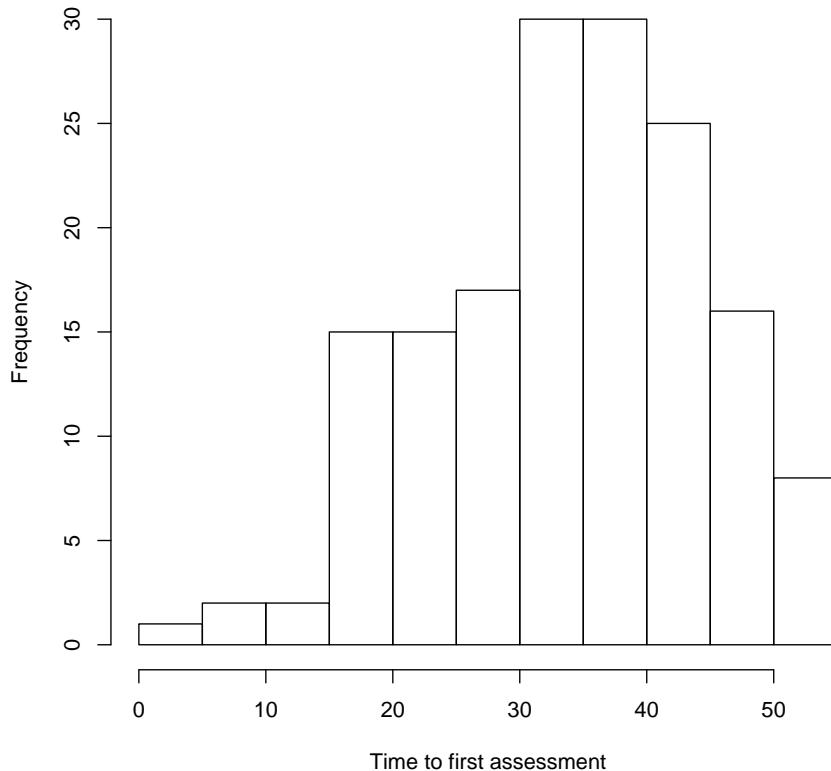
```

ref.time <- min(year.table$Year.of.first.stock.assessment, na.rm = T) - 1

year.table <- year.table %>% mutate(time = Year.of.first.stock.assessment -
  ref.time, survey.abs = Year.of.first.fishery.independent.surveys)

hist(year.table$time, 10, main = "", xlab = "Time to first assessment")

```



Get covariates from TS; looking at mean and sum @ first assessment for landings.

```

# convenience
na.fun <- function(f, ...) {
  g <- function(x) f(x, na.rm = T)
}
na.mean <- na.fun(mean)

```

```

na.sum <- na.fun(sum)

## landings --
comm.landings.ts <- comm.landings %>% select(-Stock.name, -Other.comments, -Discards.included,
      -Data.source)
comm.landings$mean.landings <- apply(comm.landings.ts, 1, na.mean)

ref.col <- which(colnames(comm.landings.ts) == "X1959")
n.col <- ncol(comm.landings.ts) - ref.col

comm.landings$land.sum <- NA
comm.landings$landings.at.assess <- NA
for (i in 1:nrow(comm.landings)) {
  asses.col <- ref.col + ifelse(!is.na(year.table$time[i]), year.table$time[i],
    n.col)
  comm.landings$land.sum[i] <- na.sum(comm.landings.ts[i, 1:asses.col])
  comm.landings$landings.at.assess[i] <- comm.landings.ts[i, asses.col]
}

```

Same for Rec landings, although I am not sure if these are useable since many rows are NAs.

```

## landings --
rec.landings.ts <- rec.landings %>% select(-Stock.name, -Discards.included,
      -Data.source)
rec.landings$mean.landings <- apply(rec.landings.ts, 1, na.mean)
rec.landings$landings.at.assess <- apply(rec.landings.ts, 1, na.mean)

ref.col <- which(colnames(rec.landings.ts) == "X1959")
n.col <- ncol(rec.landings.ts) - ref.col

rec.landings$land.sum <- NA
rec.landings$landings.at.assess <- NA
for (i in 1:nrow(rec.landings)) {
  asses.col <- ref.col + ifelse(!is.na(year.table$time[i]), year.table$time[i],
    n.col)
  rec.landings$land.sum[i] <- na.sum(rec.landings.ts[i, 1:asses.col])
  rec.landings$landings.at.assess[i] <- rec.landings.ts[i, asses.col]
}

rec.landings$land.sum[rec.landings$land.sum == 0] <- NA

```

Same for Price;

```

## landings --
price.ts <- price %.% select(-Stock.name, -Region.used.for.price.data, -Other.comments)
price$mean.price <- apply(price.ts, 1, na.mean)

ref.col <- which(colnames(price.ts) == "X1959")
n.col <- ncol(price.ts) - ref.col

price$price.sum <- NA
price$price.at.assess <- NA
for (i in 1:nrow(price)) {
  asses.col <- ref.col + ifelse(!is.na(year.table$time[i]), year.table$time[i],
  n.col)
  price$price.sum[i] <- na.sum(price.ts[i, 1:asses.col])
  price$price.at.assess[i] <- price.ts[i, asses.col]
}
price$price.sum[price$price.sum == 0] <- NA

year.table$Rebuild <- ordered(year.table$Rebuilding.plan.history, levels = c("never",
  "previously", "currently"))

with(year.table, table(Rebuild))

## Rebuild
##      never previously currently
##      106          13         25

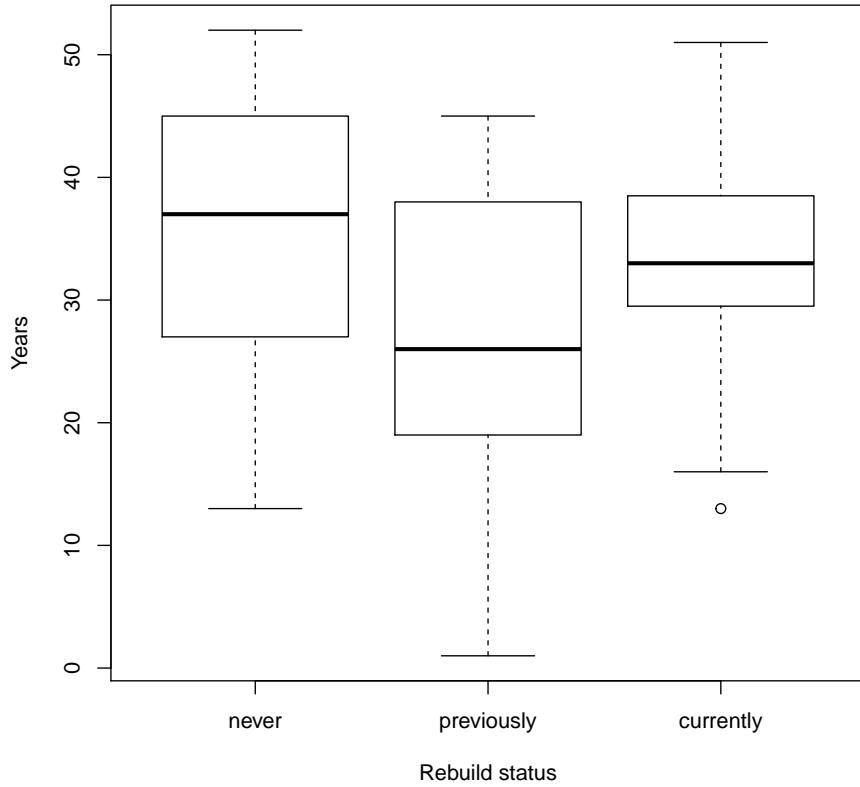
with(year.table, table(Region))

## Region
## Alaska    USEC    USNE    USSE    USWC
##      51        9       49       34       50

# set north-east as reference treatment
RC <- contr.treatment(levels(factor(year.table$Region)), base = 3)

boxplot(time ~ Rebuild, data = year.table, xlab = "Rebuild status", ylab = "Years")

```



## 2 Simple Bayesian Weibull survival model

Try a Bayesian truncated Weibull model to keep it simple to start with:

```
# subset to data with price and landings data

model.data <- year.table %>% mutate(price = price$mean.price, landings = comm.landings$mean
  filter(!is.na(price), !is.na(landings))

# assessment time
a.time <- model.data$time

# true/false censoring
censored <- as.numeric(is.na(a.time))
```

```

# censor time - improve here from the arbitrary 2010 cutoff for censored
# (non-assessed stocks)
ctime <- a.time
ctime[is.na(a.time)] <- 2010 - ref.time

# initial values for censored observations
time.inits <- ctime + 1
time.inits[!is.na(a.time)] <- NA

# habitat and family random effect - note - taxon is half way between
# habitat and family variables, try that too sometime

afs <- function(x) as.numeric(as.factor(x))

hab <- with(model.data, afs(habitat_MM))
n.hab <- length(unique(hab))

fam <- with(model.data, afs(Family))
n.fam <- length(unique(fam))

# fixed effect for regions
region <- data.frame(with(model.data, model.matrix(~Region, contrasts.arg = list(Region = R
# 

# Covariate dataframe
covs <- model.data %.% select(Lmax..cm., TL, Recreational.pc.catch, Year.of.fishery.developm
    price, landings)

# replace TL for CA spiny lobster with something approximate for now since I
# can't find a good value
covs$TL[is.na(covs$TL)] <- 3.2

# scale covariates for comparison
sc.covs <- data.frame(apply(covs, 2, function(x) (x - mean(x))/(2 * sd(x))))
COVS <- cbind(region, sc.covs)
n.covs <- ncol(COVS)
n.stocks <- nrow(COVS)

# set up jags model

require(rjags)

## Loading required package: rjags
## Loading required package: coda
## Loading required package: lattice

```

```

## Linked to JAGS 3.3.0
## Loaded modules: basemod, bugs

jags.data <- list(COVS = COVS, n.covs = n.covs, n.stocks = n.stocks, hab = hab,
  fam = fam, n.hab = n.hab, n.fam = n.fam, ctime = ctime, a.time = a.time,
  censored = censored)

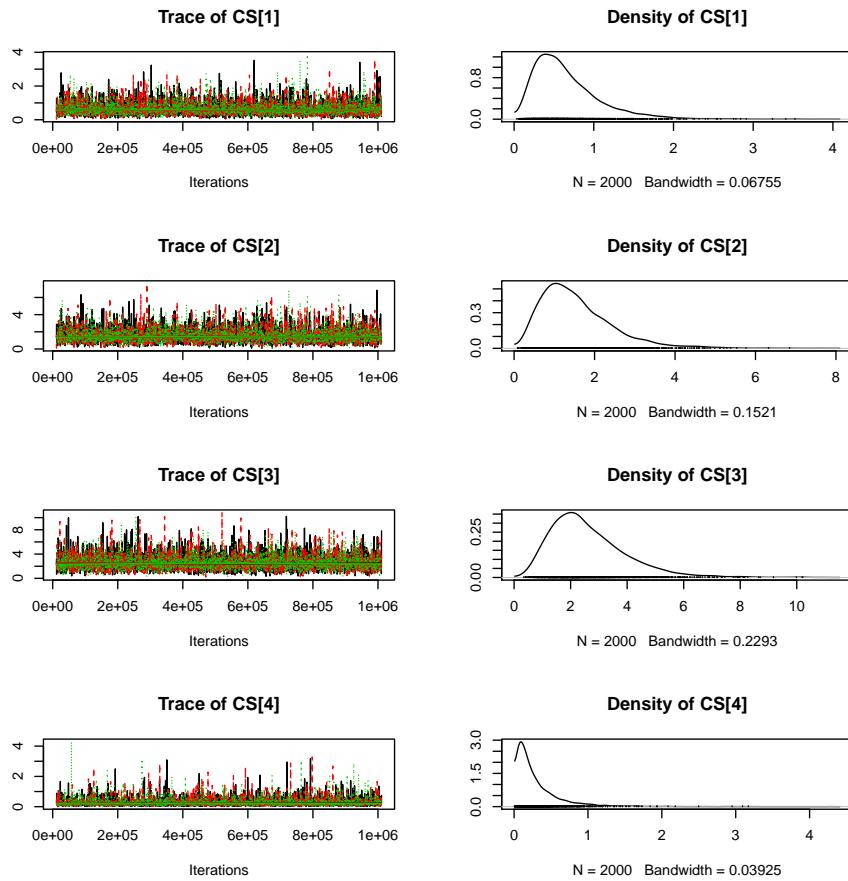
# run model - short run for now...
JM <- jags.model("Weib_surv.R", inits = list(a.time = time.inits), data = jags.data,
  n.chains = 3)

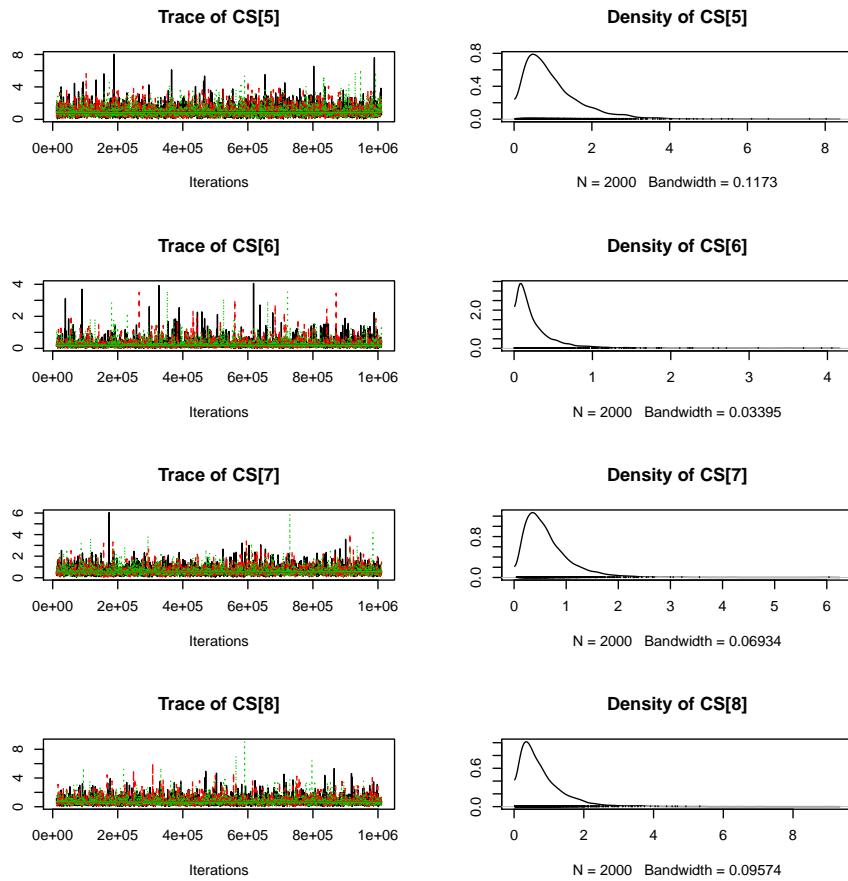
update(JM, n.iter = 10000)

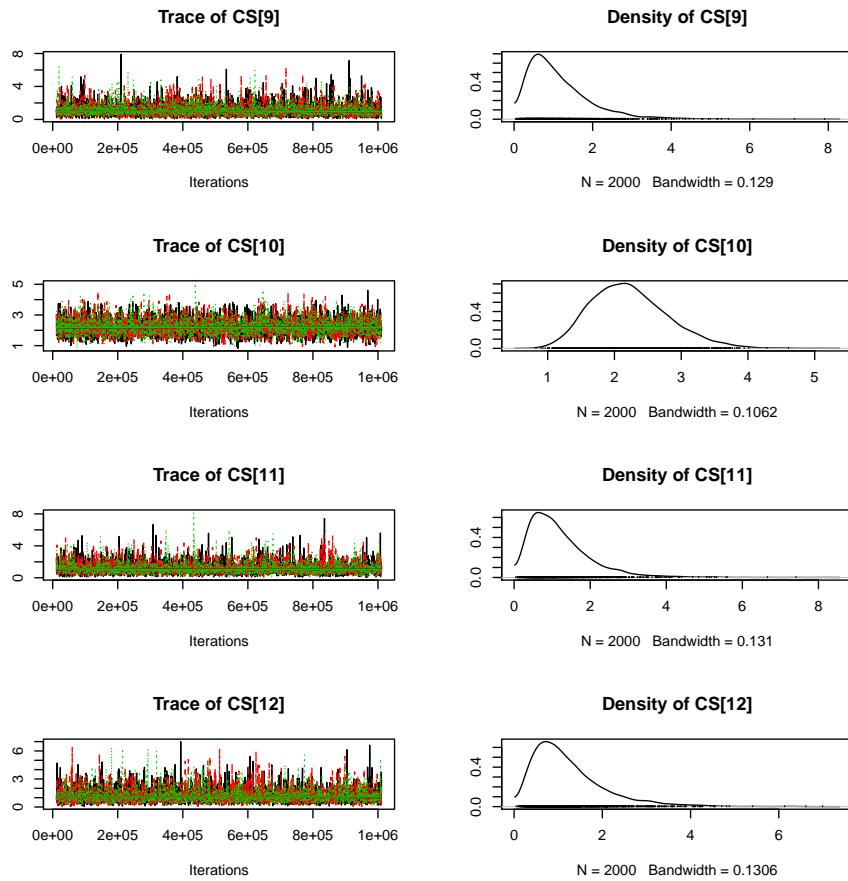
a.out <- coda.samples(JM, variable.names = c("betas", "habitat", "fp.var", "CS"),
  n.iter = 1e+06, thin = 500)

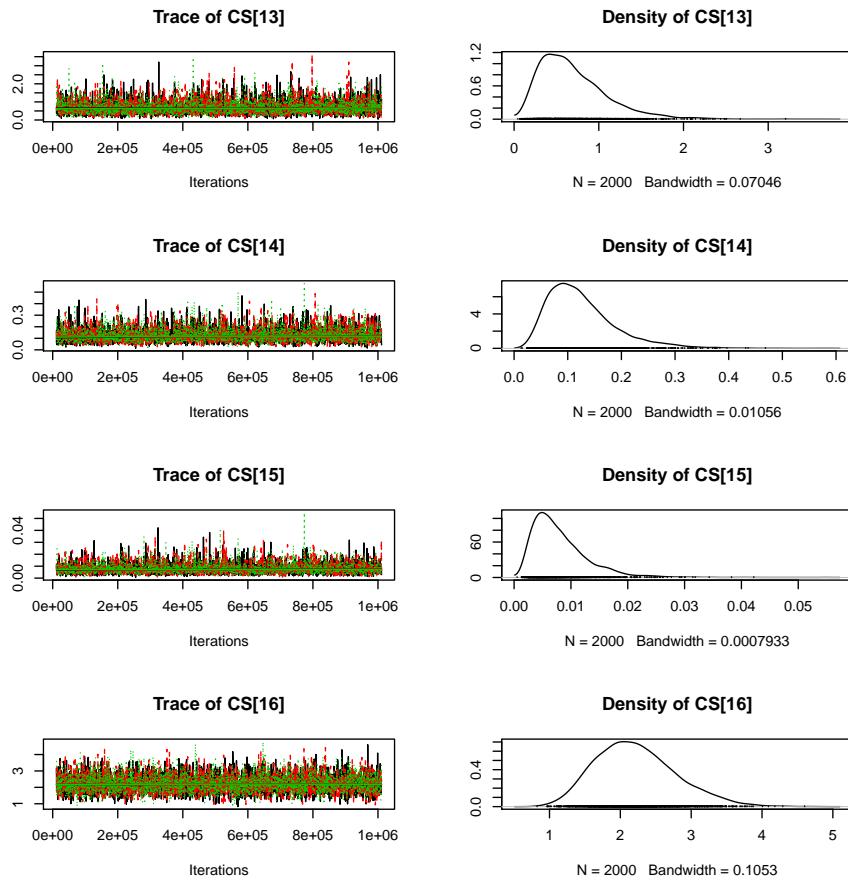
plot(a.out)

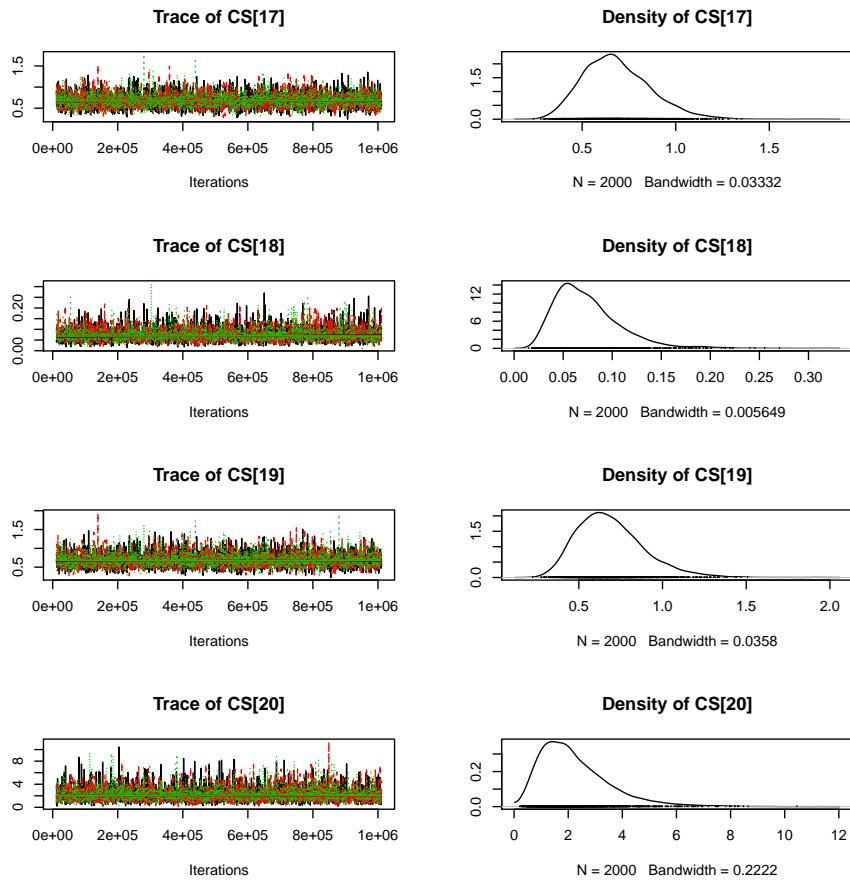
```

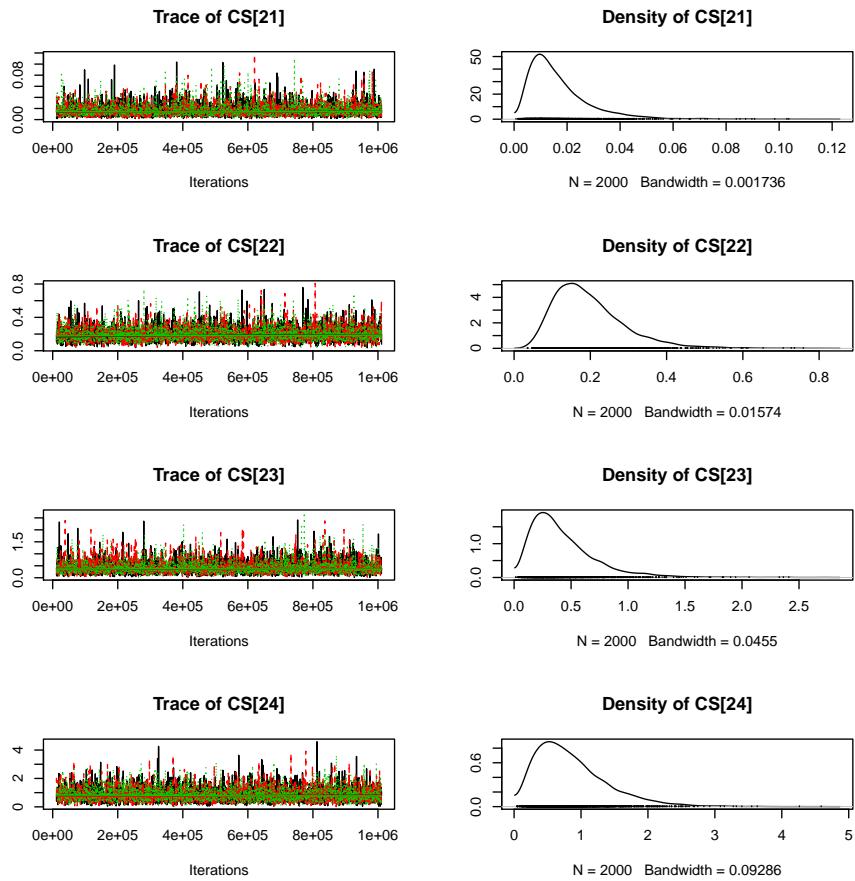


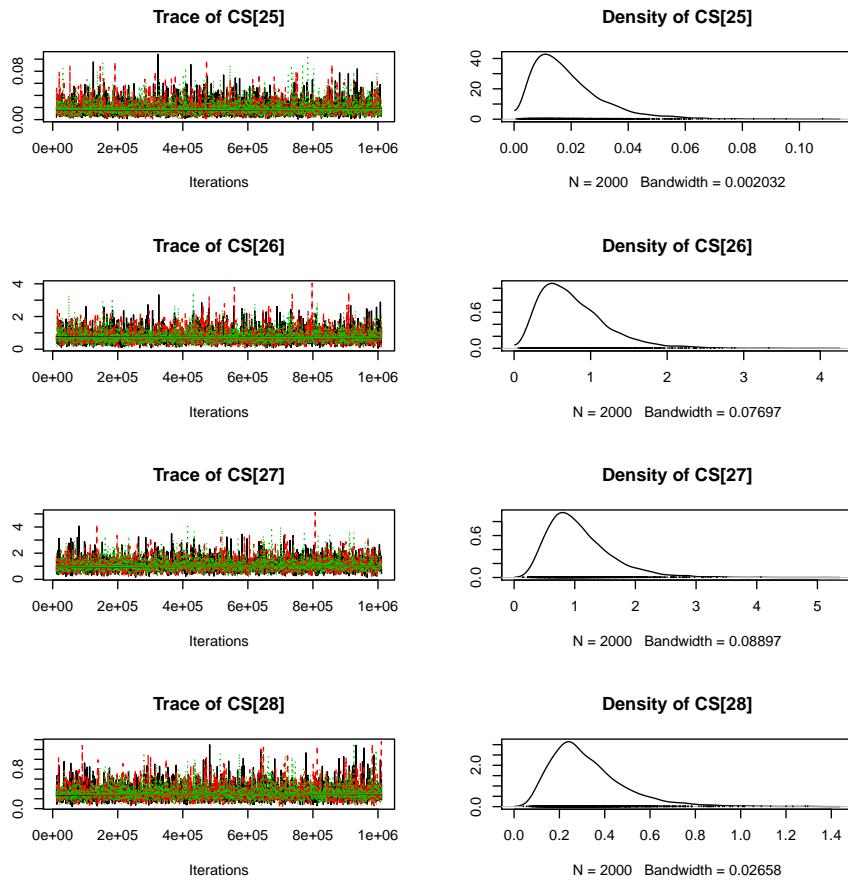


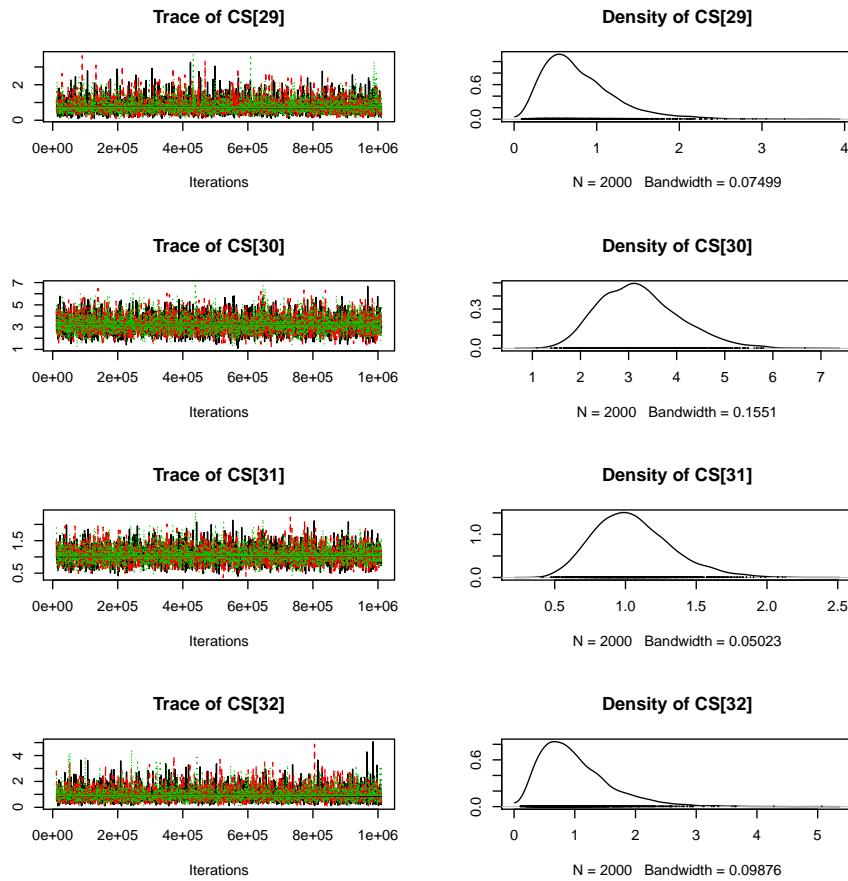


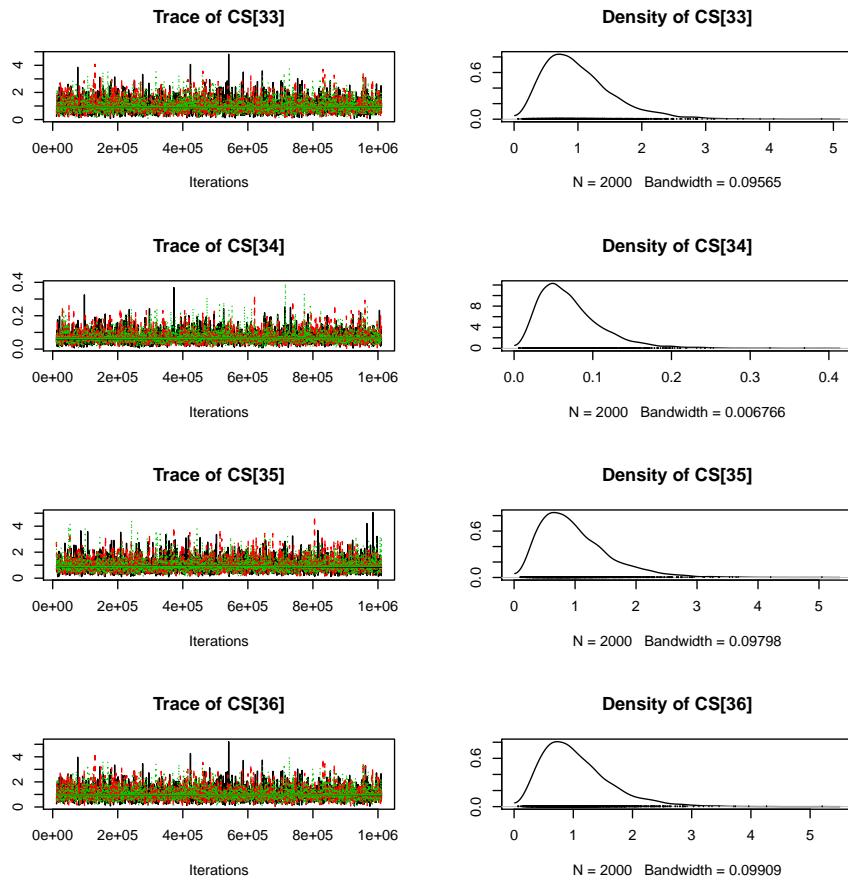


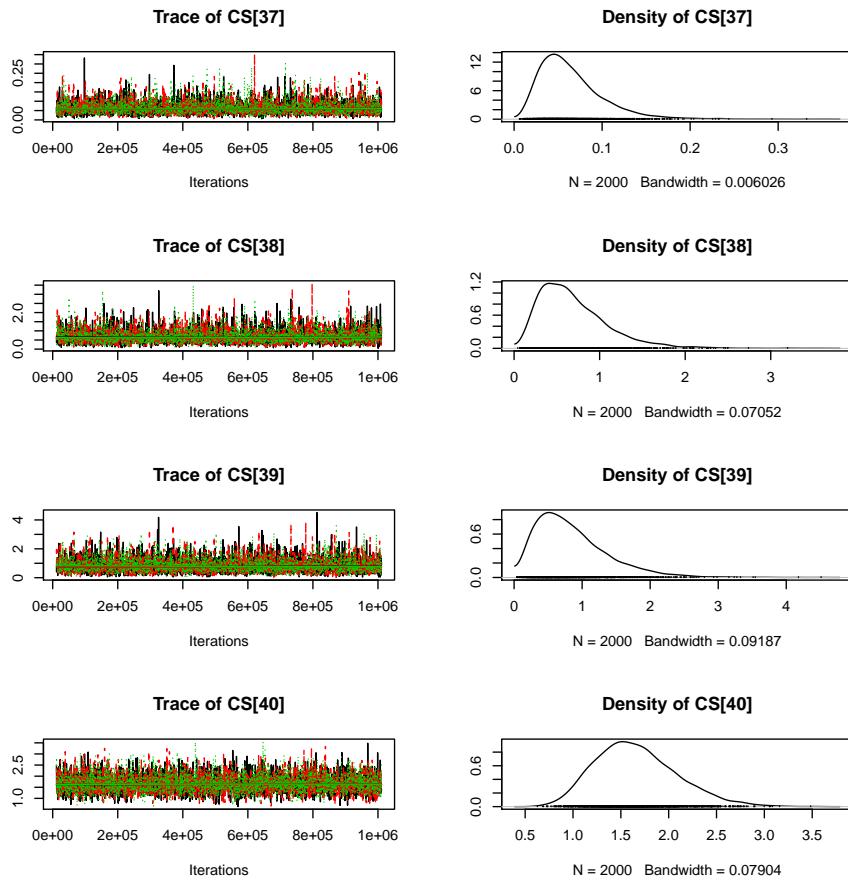


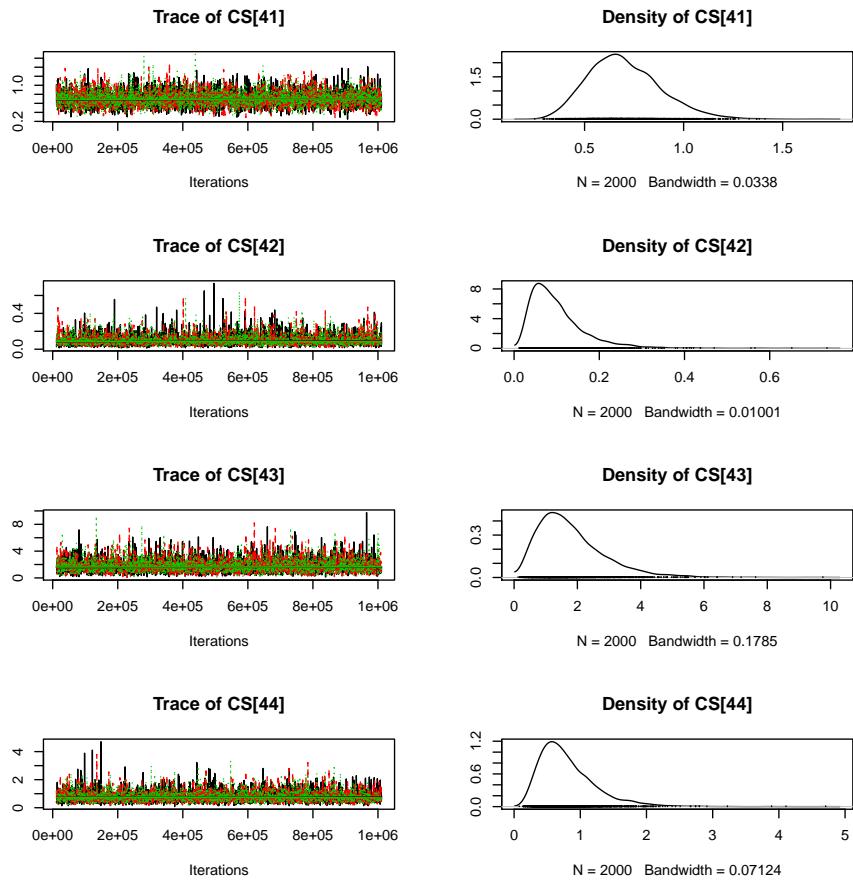


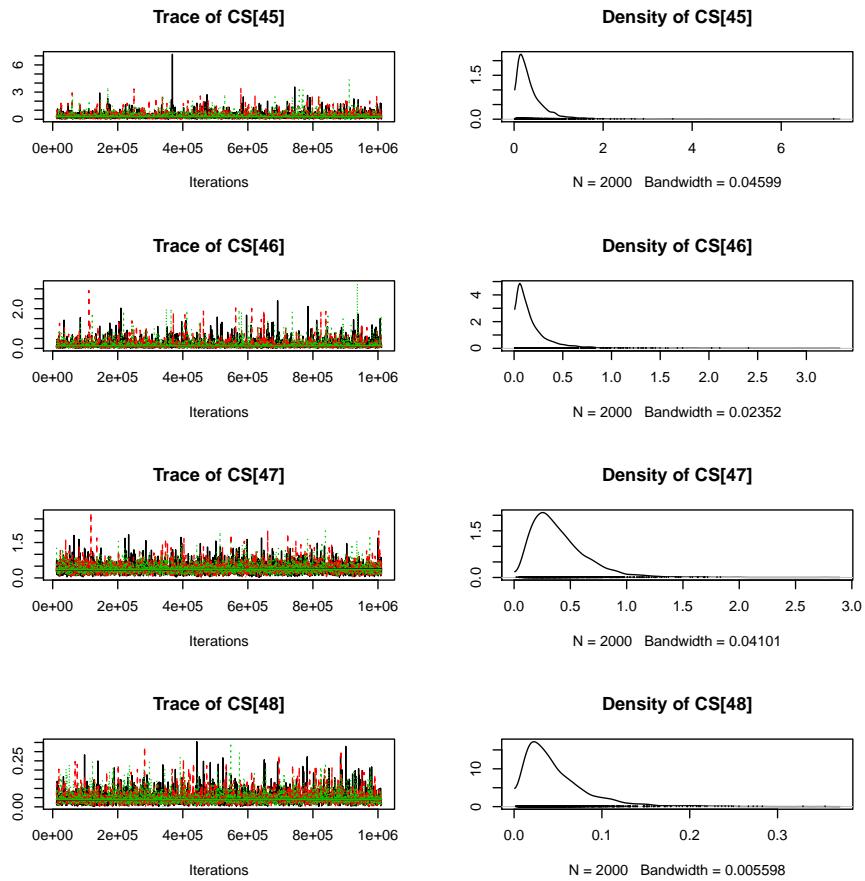


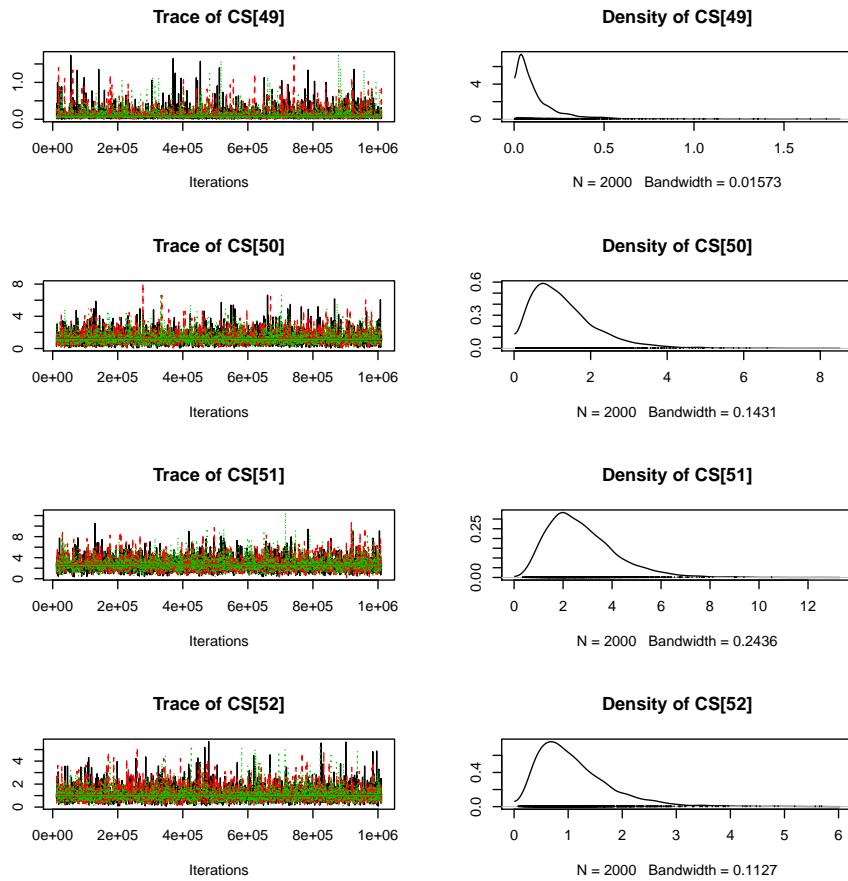


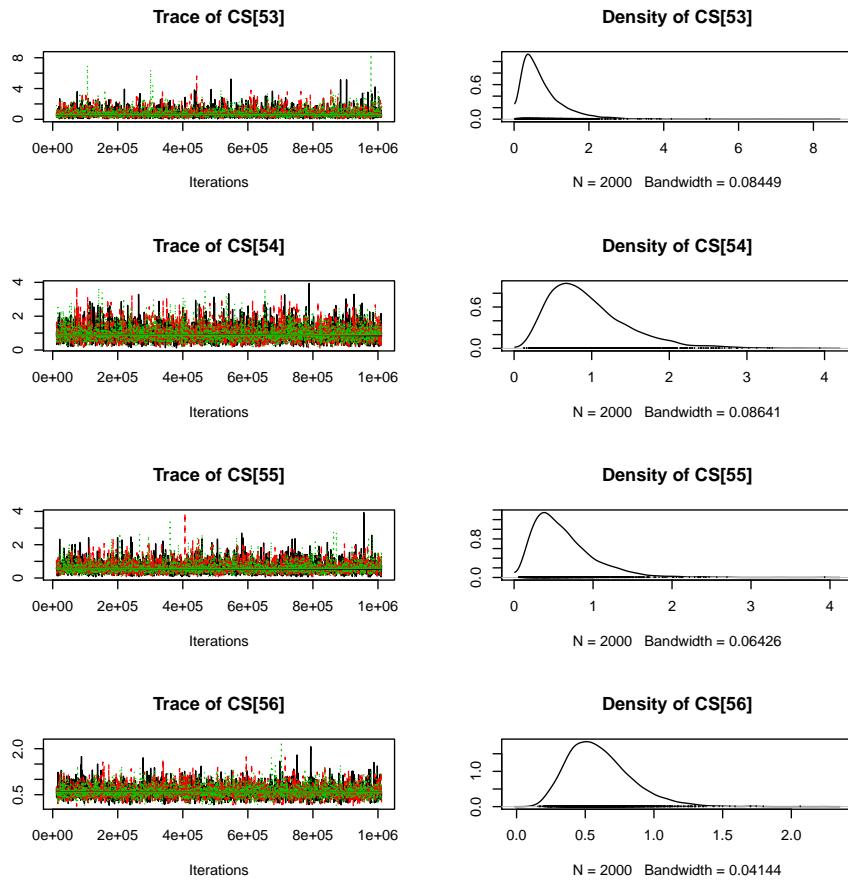


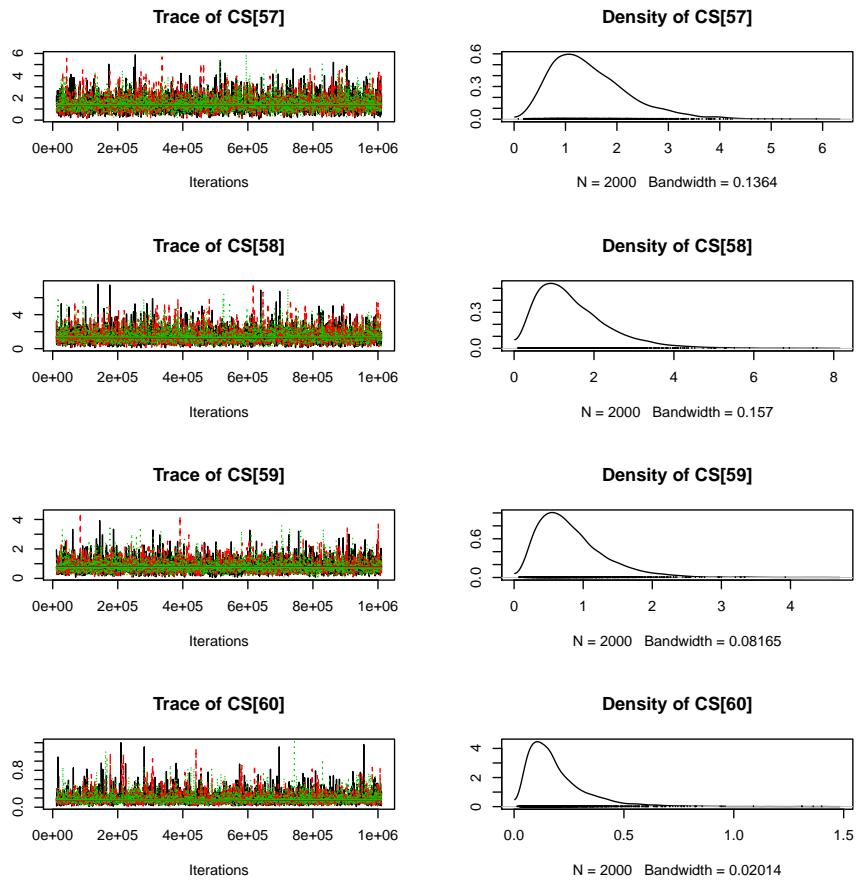


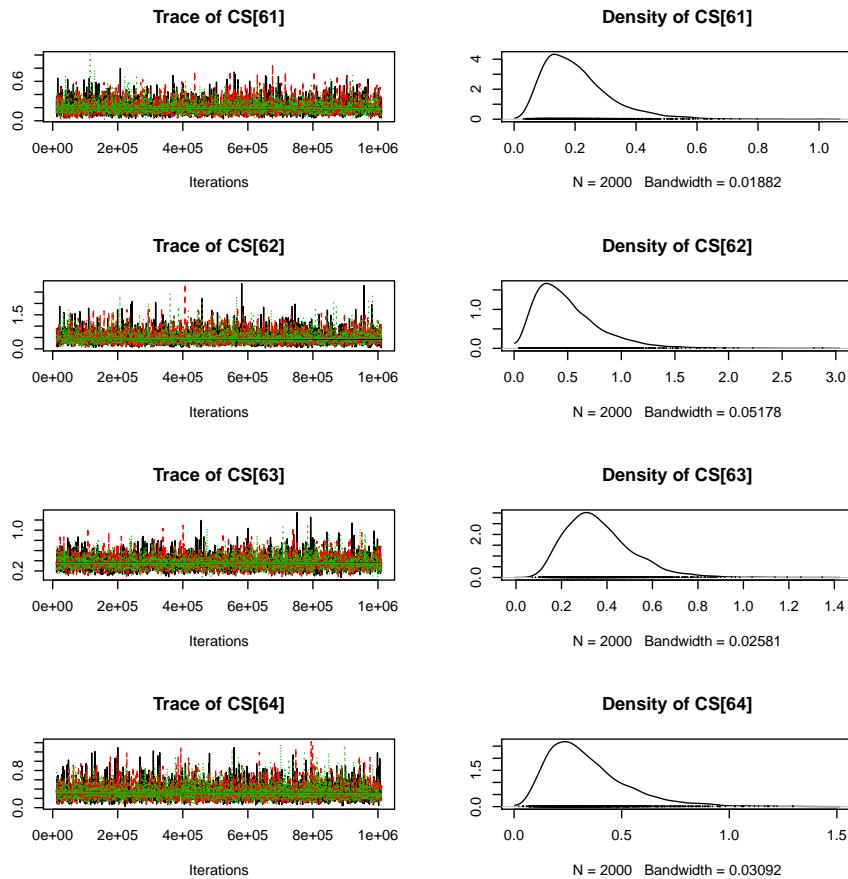


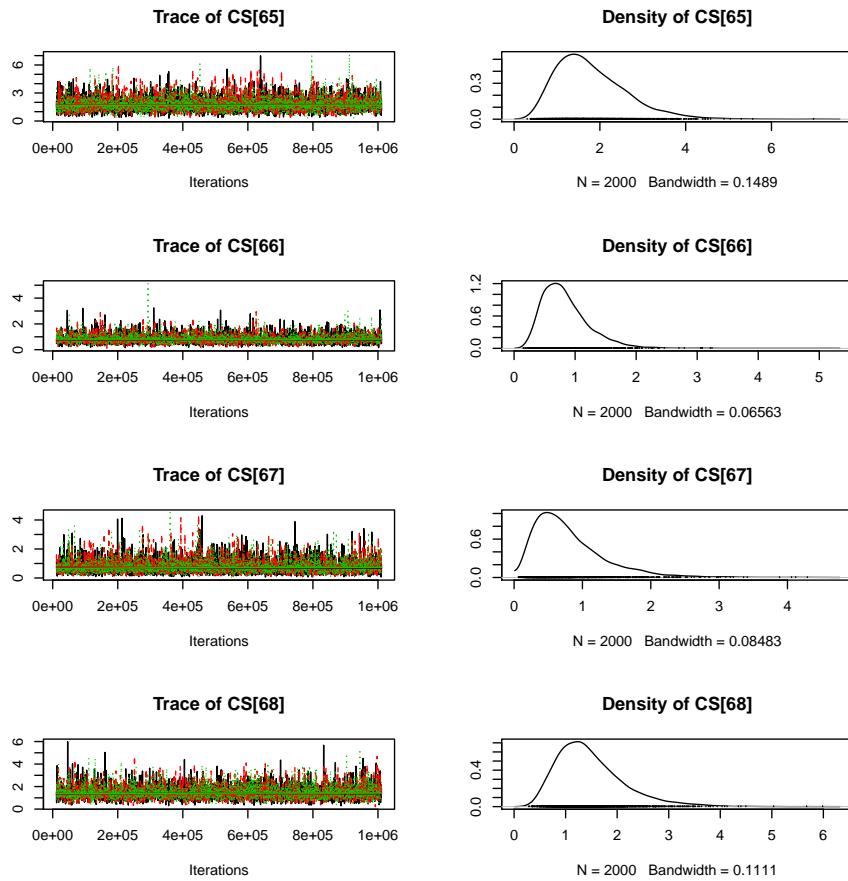


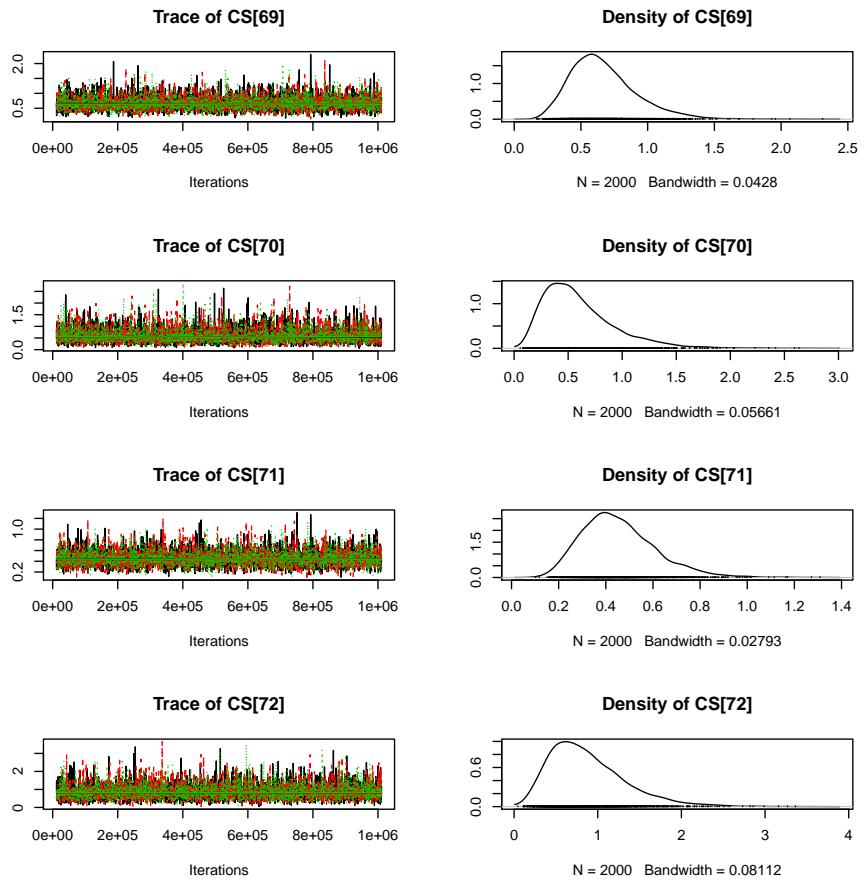


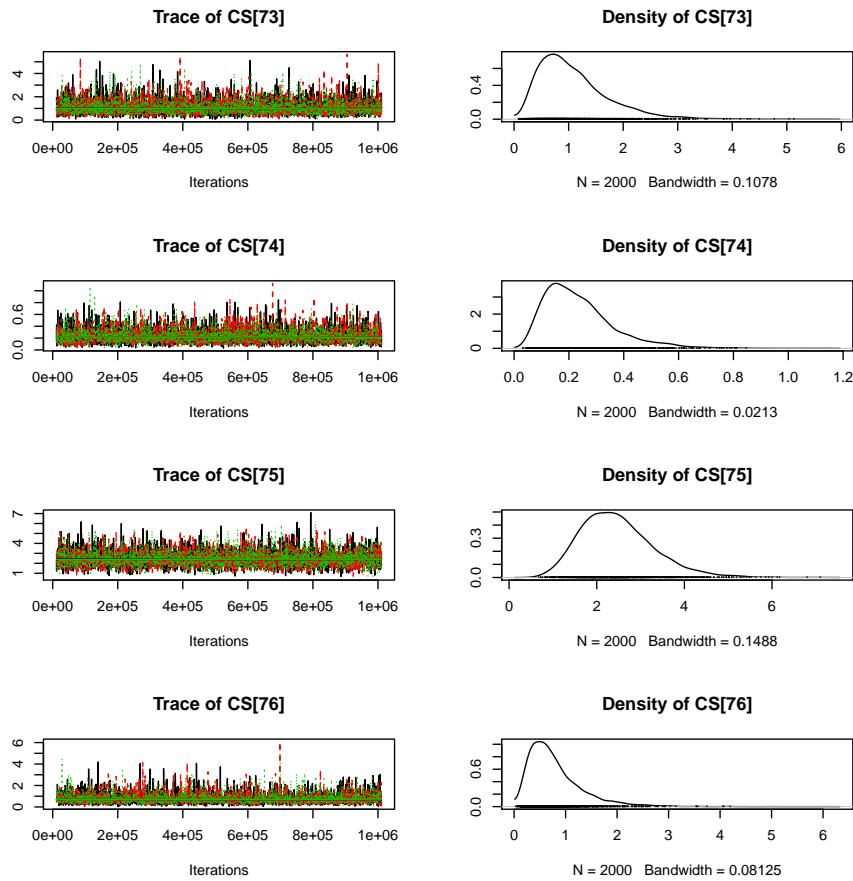


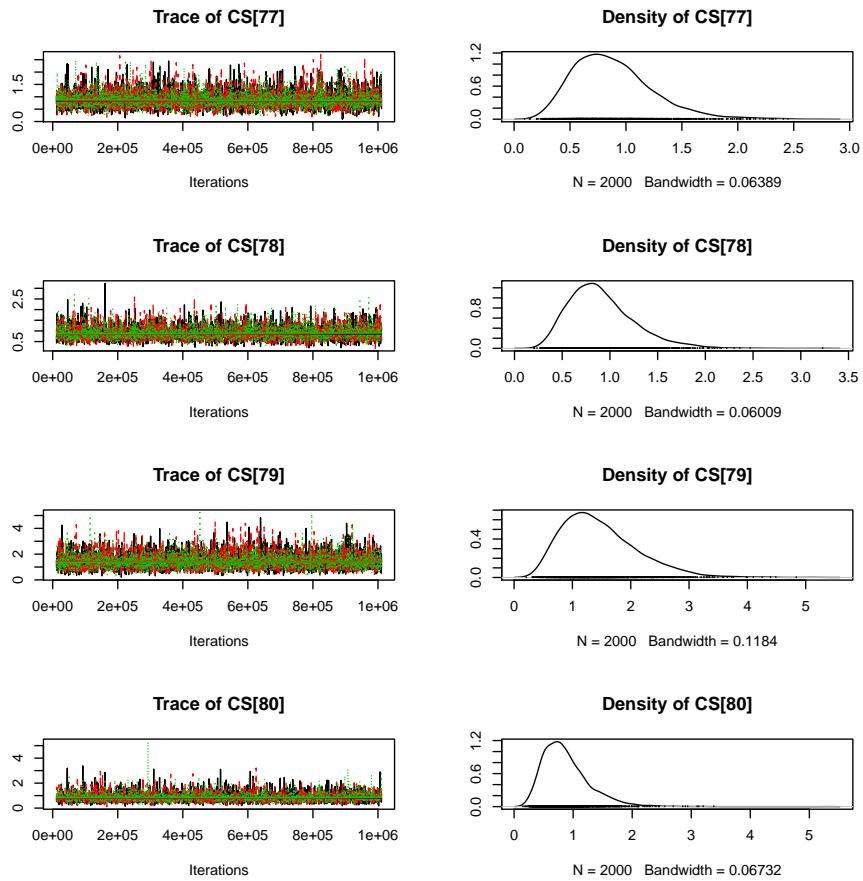


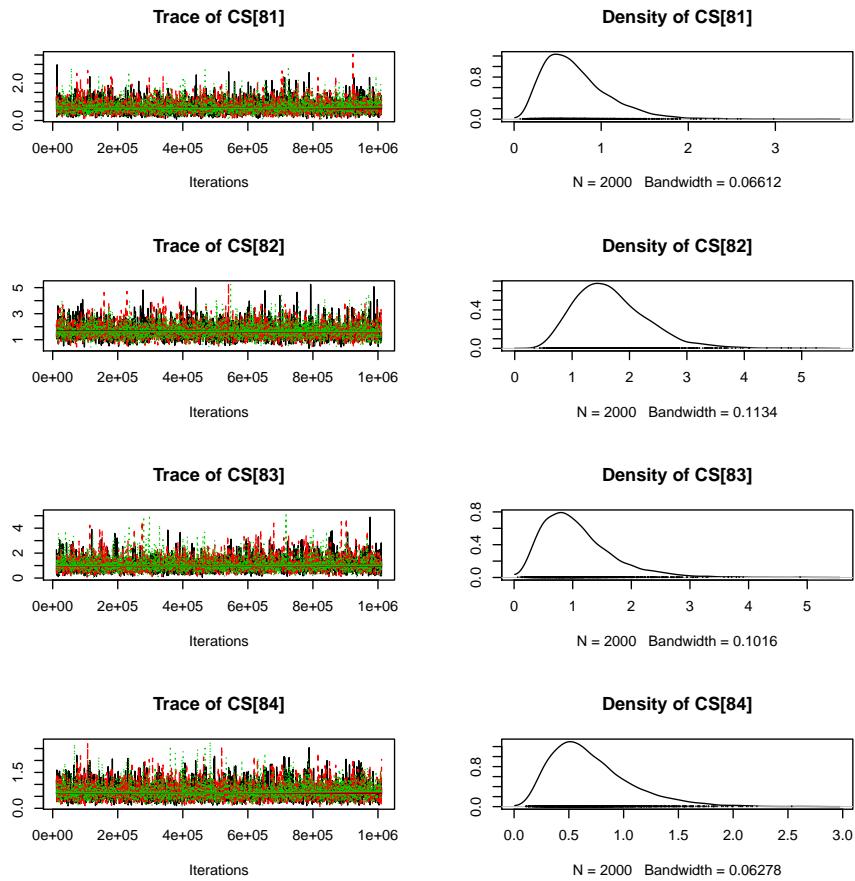


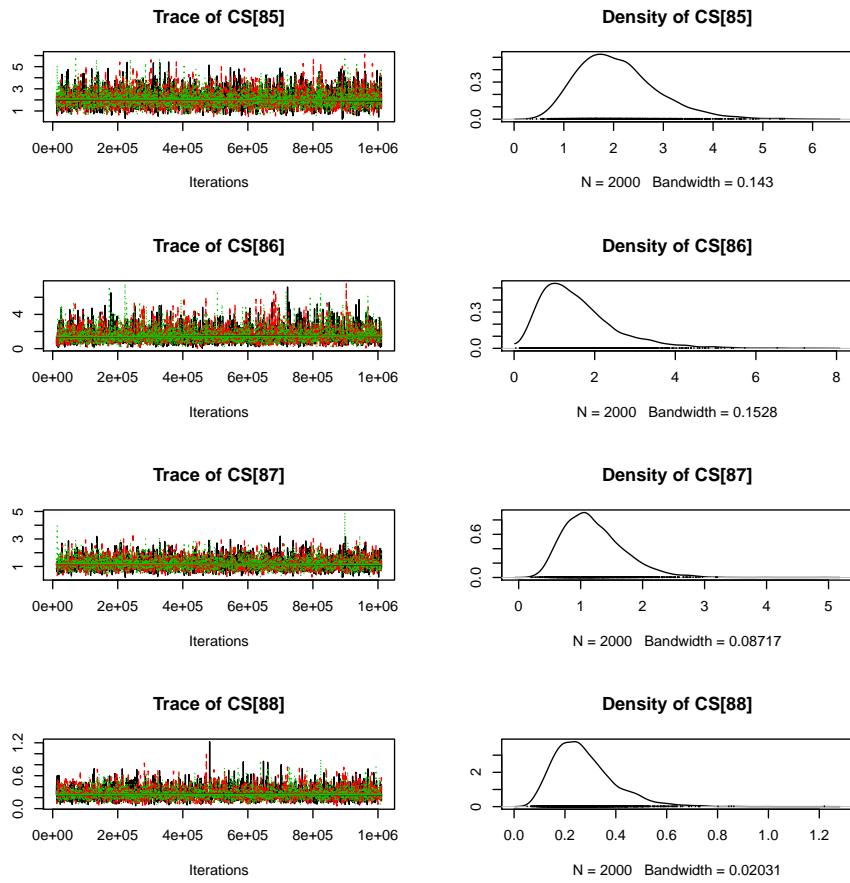


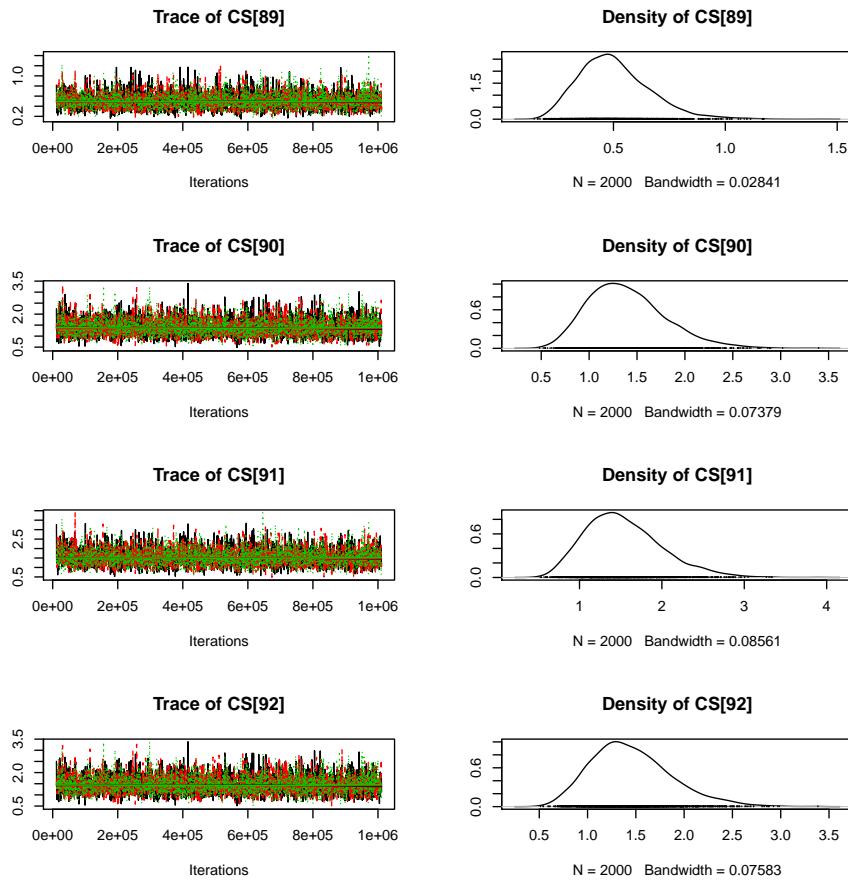


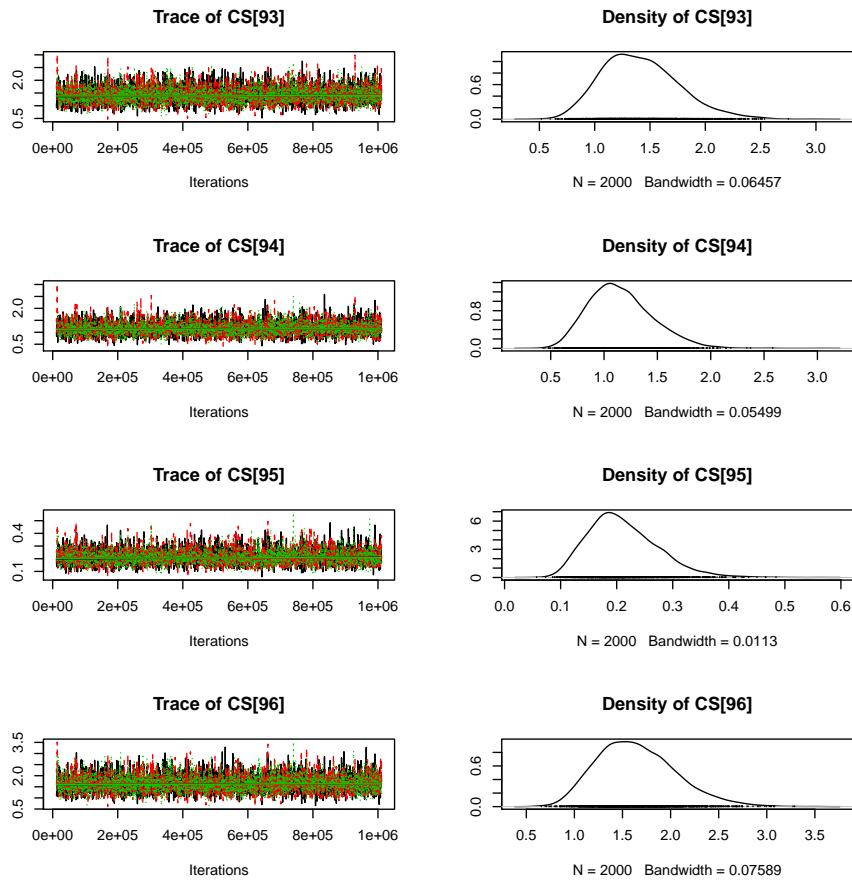


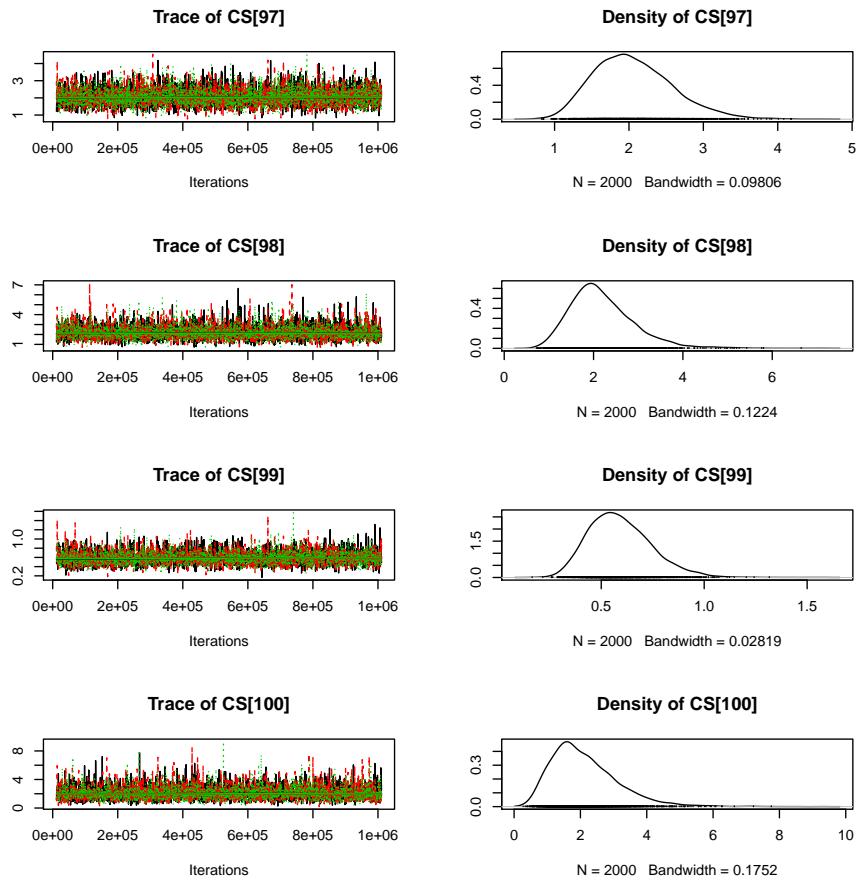


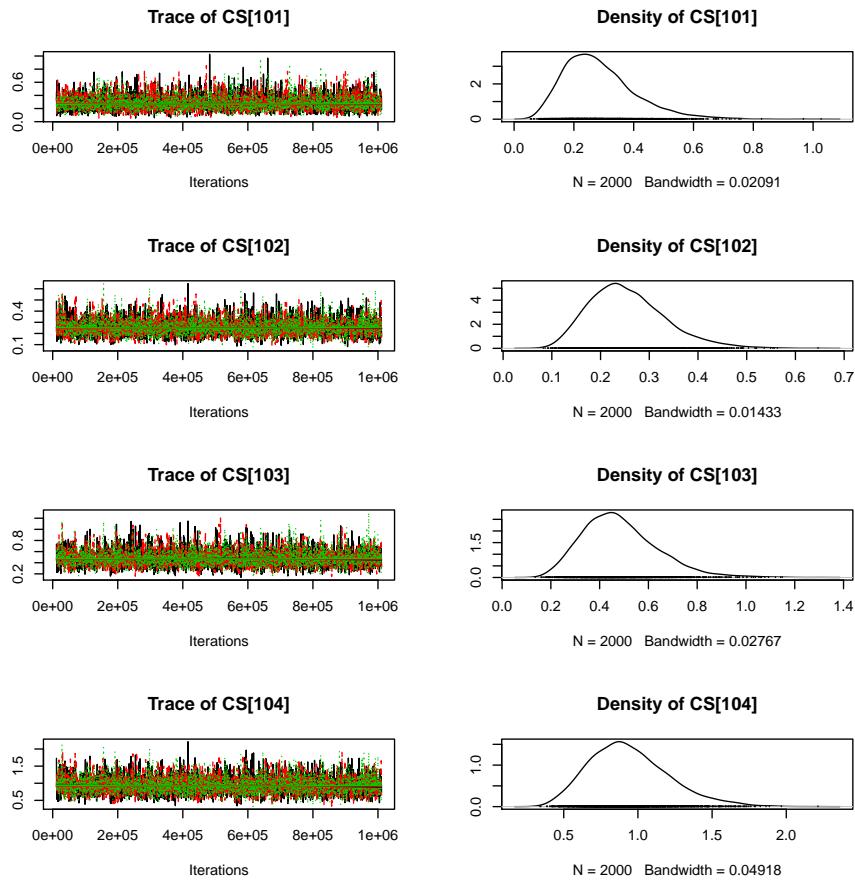


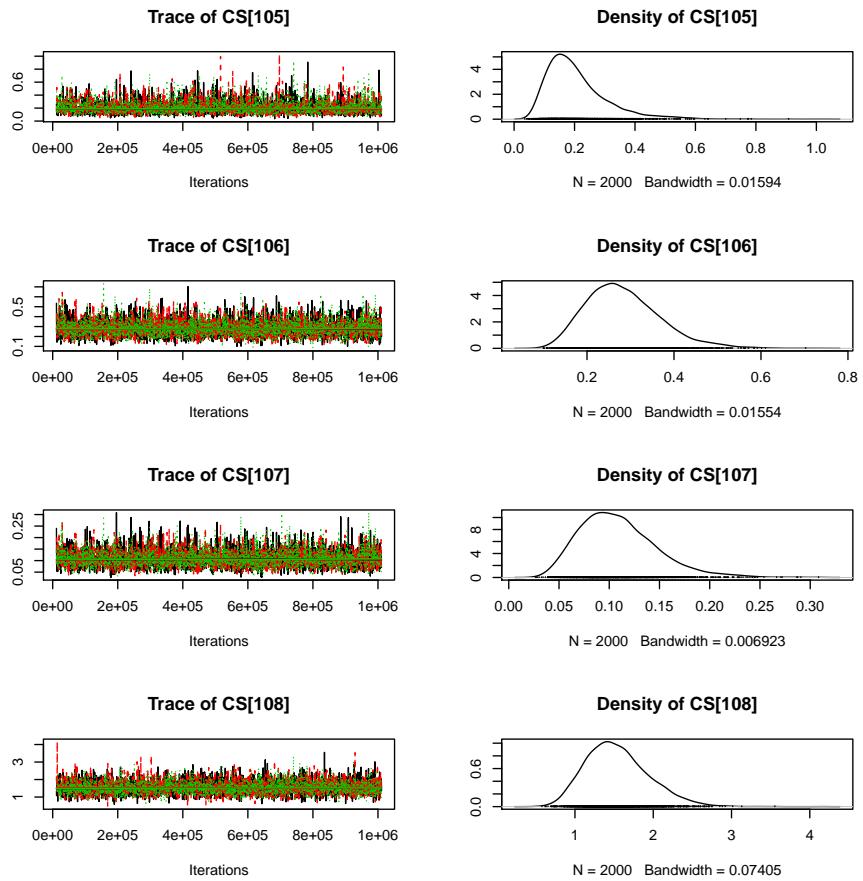


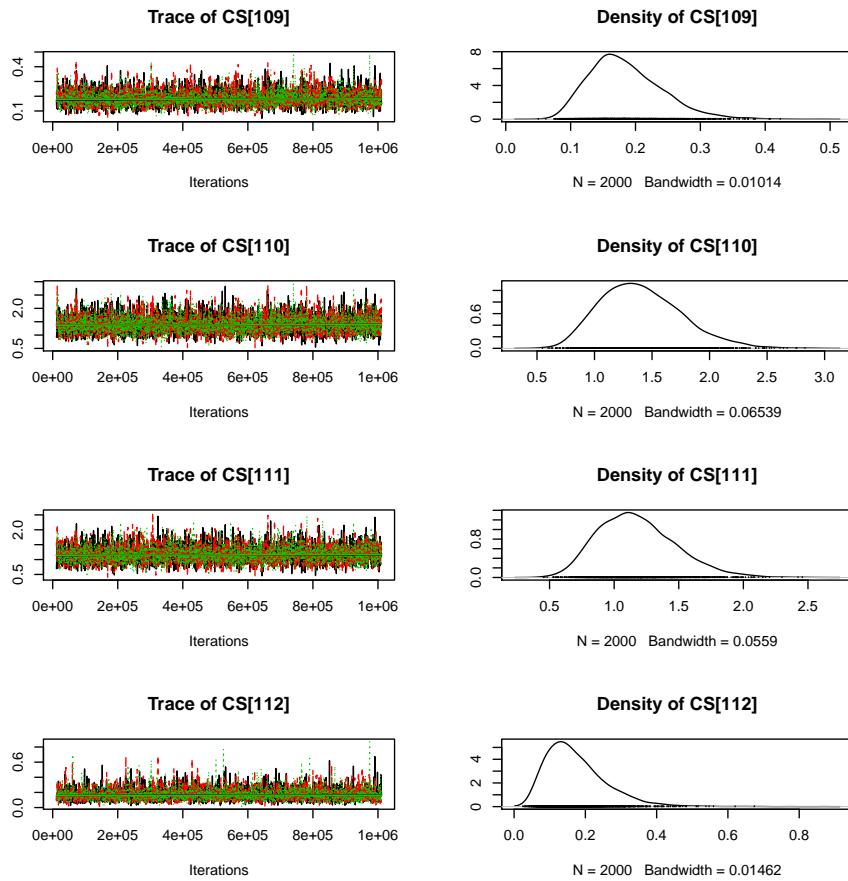


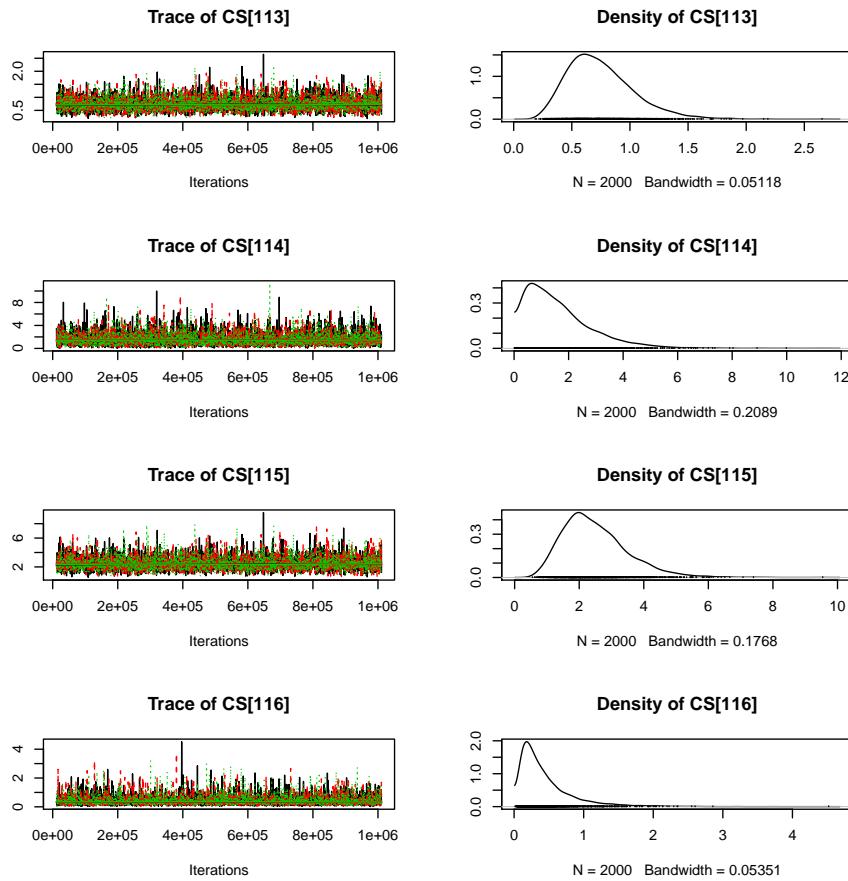


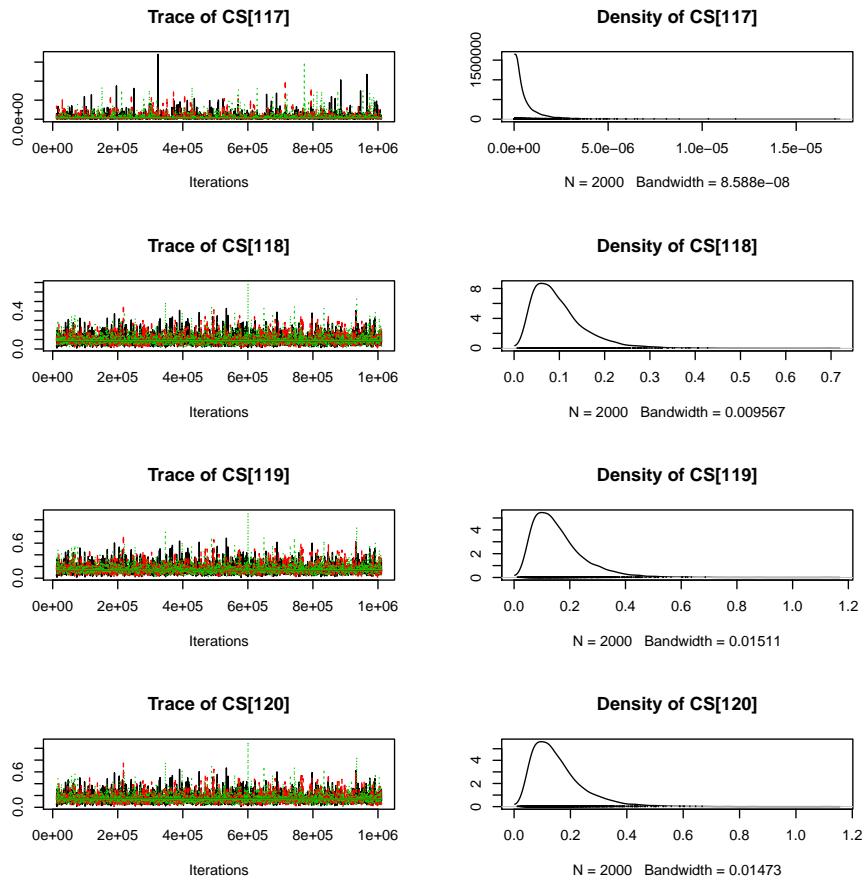


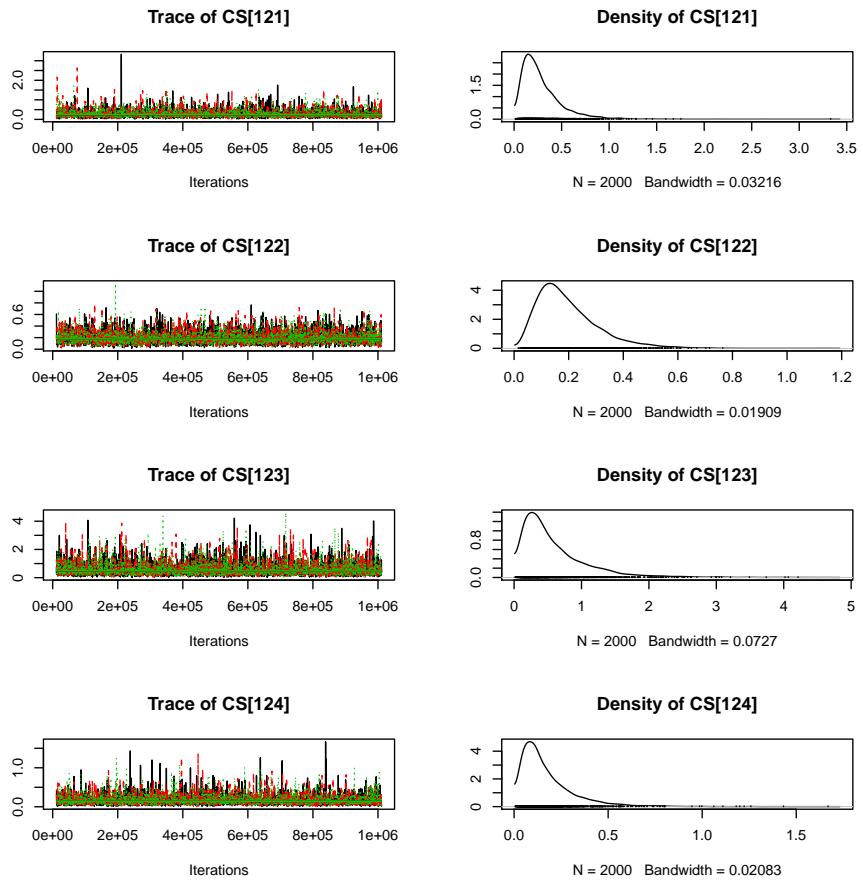


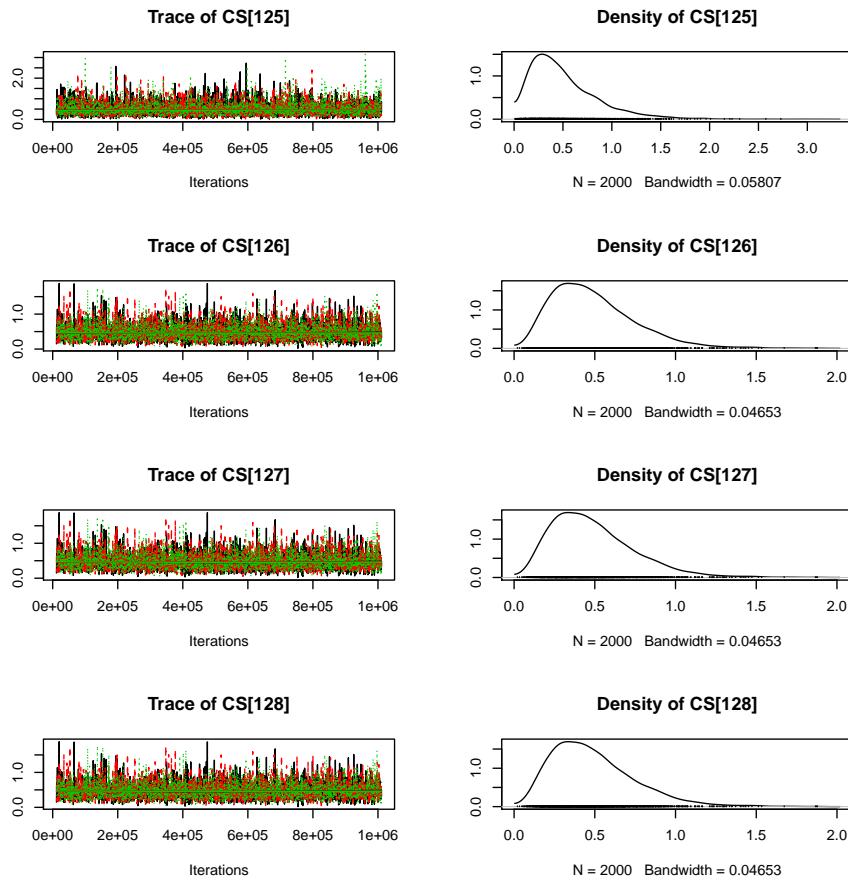


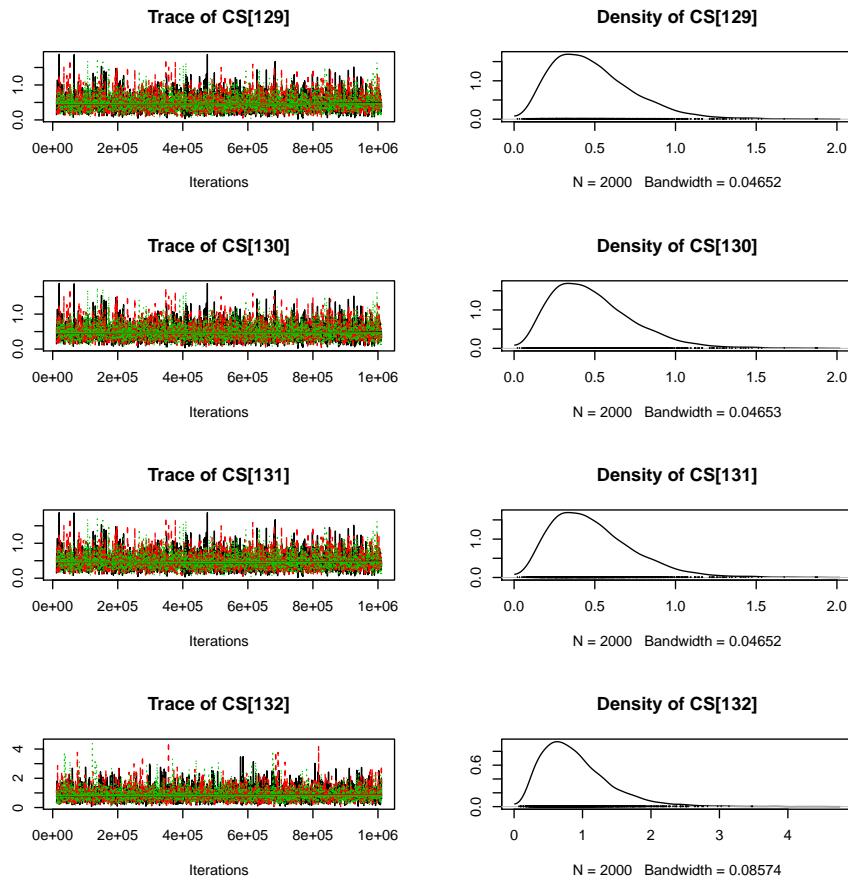


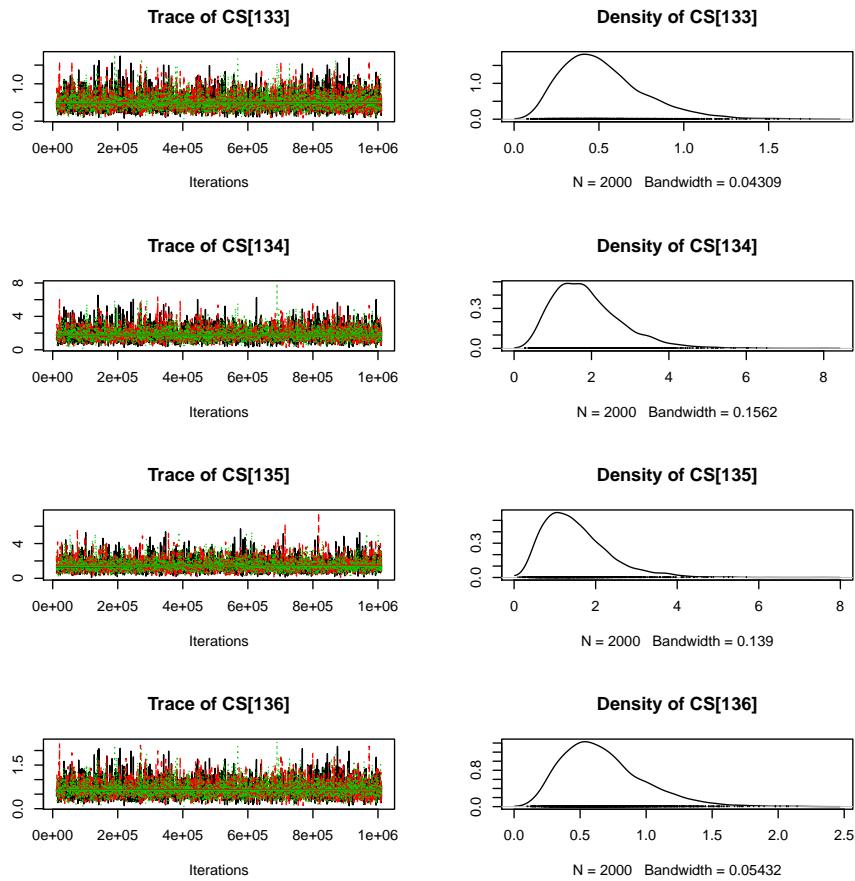


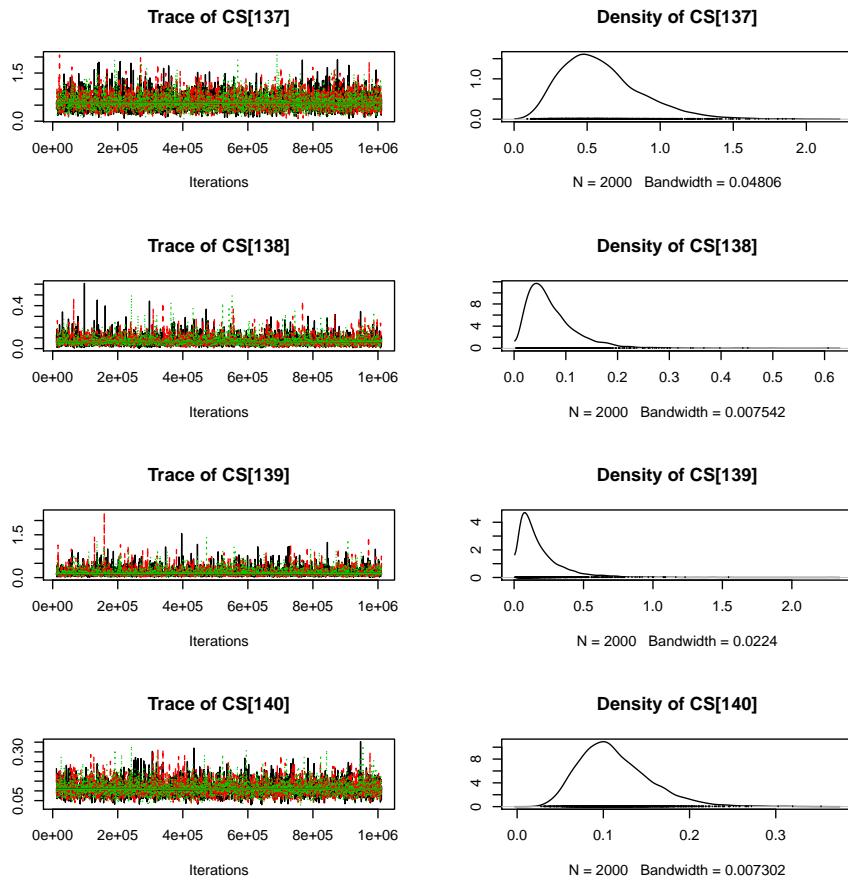


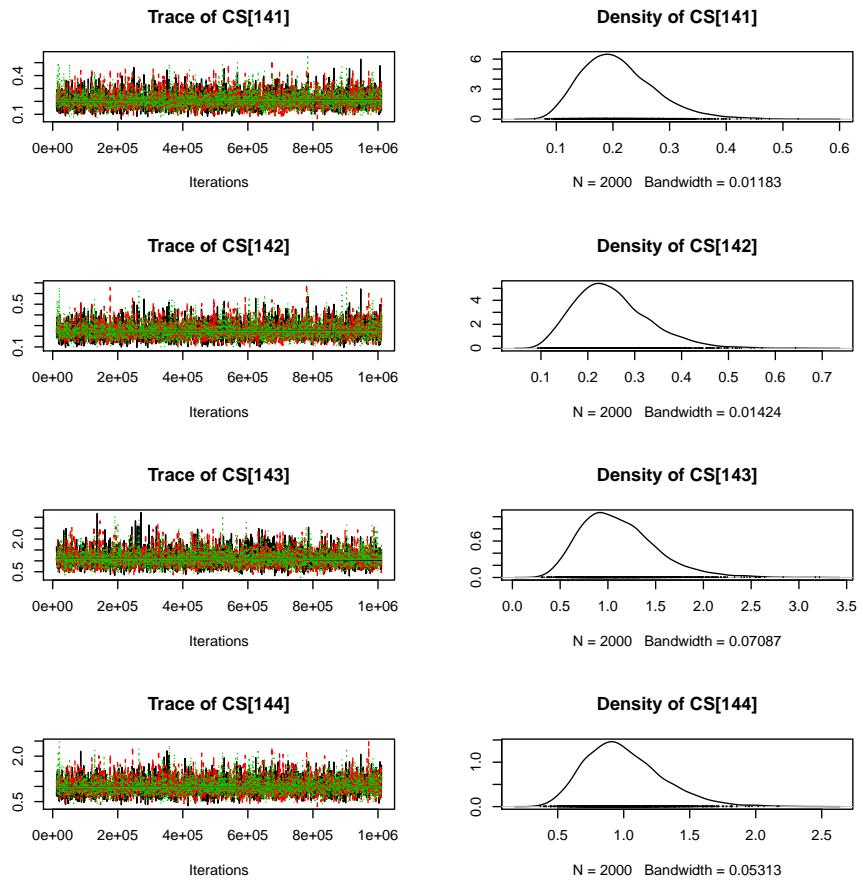


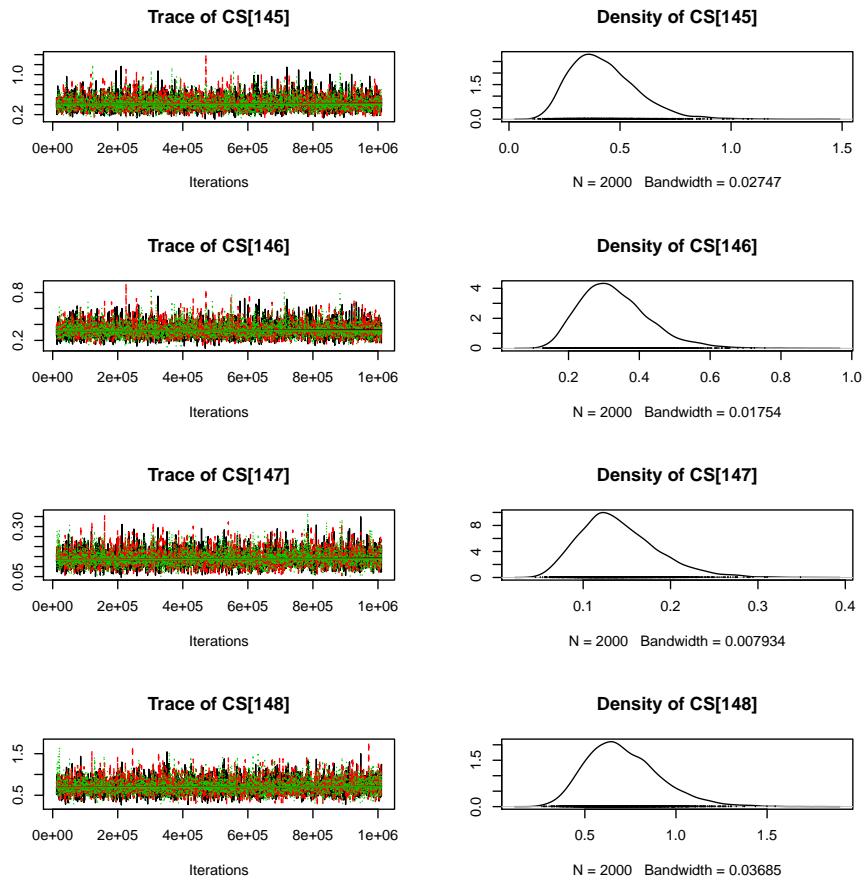


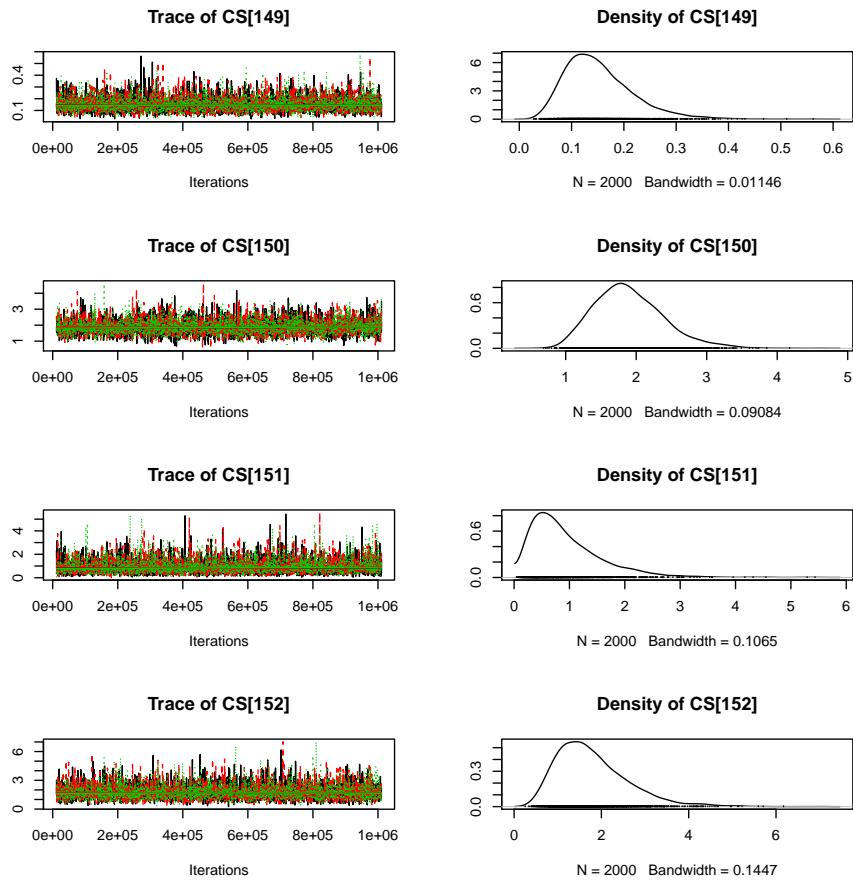


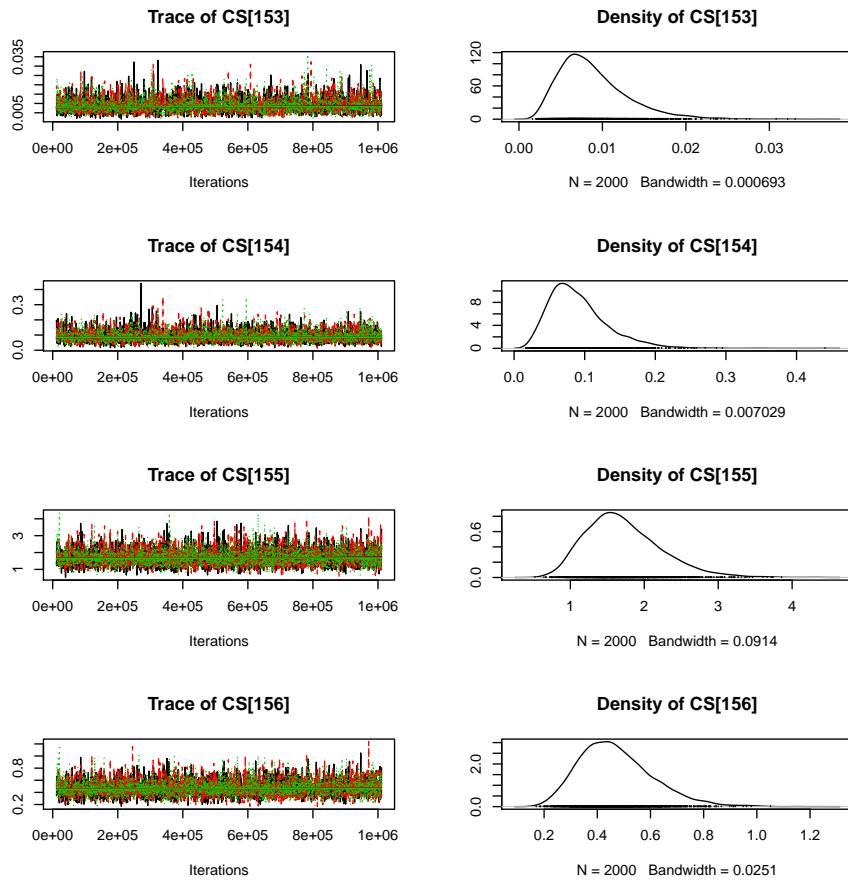


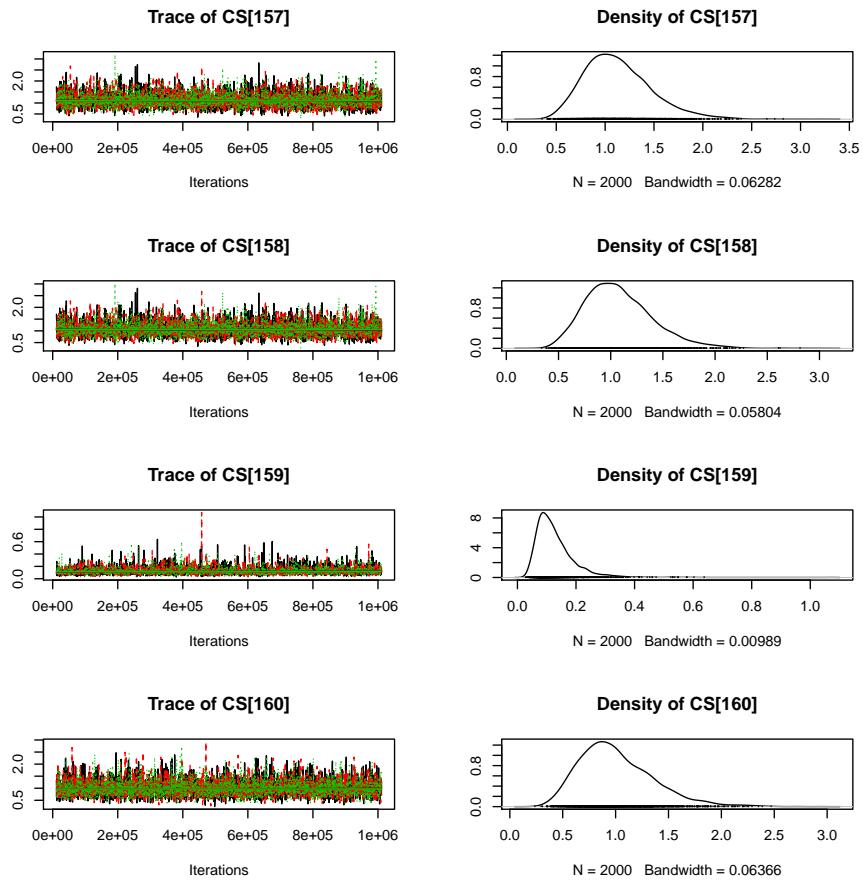


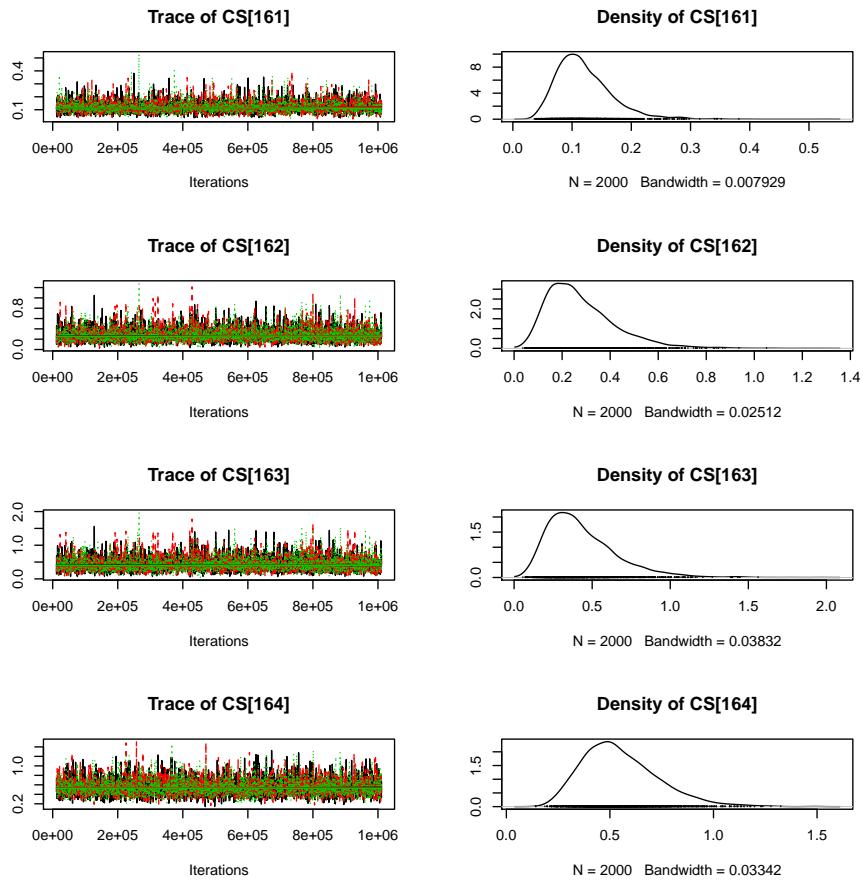


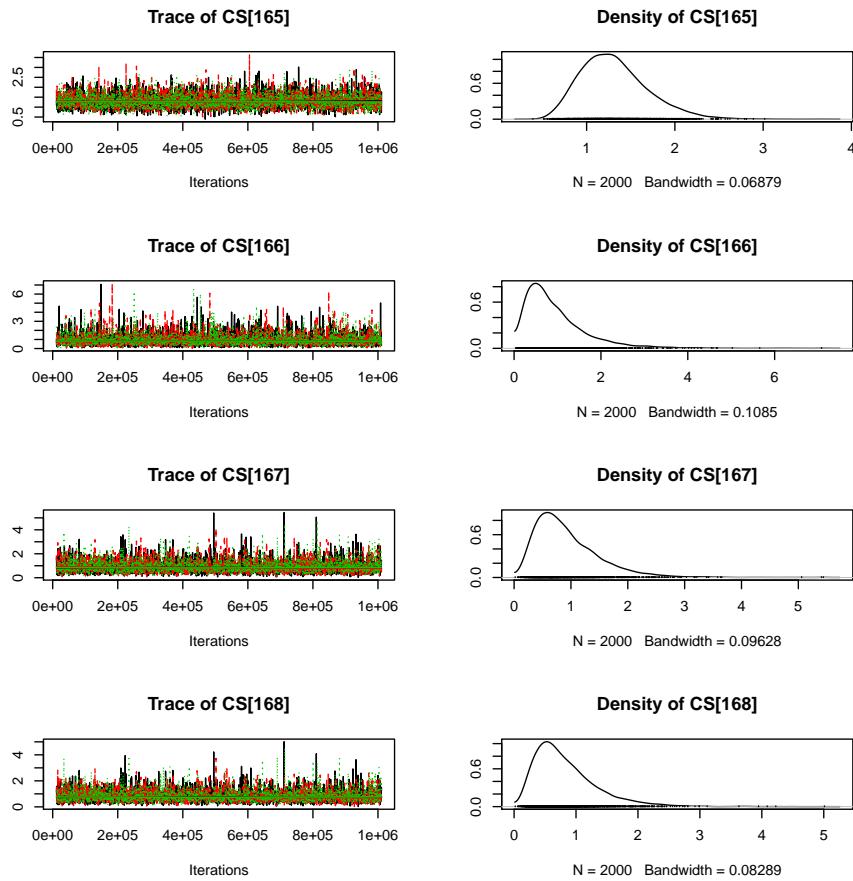


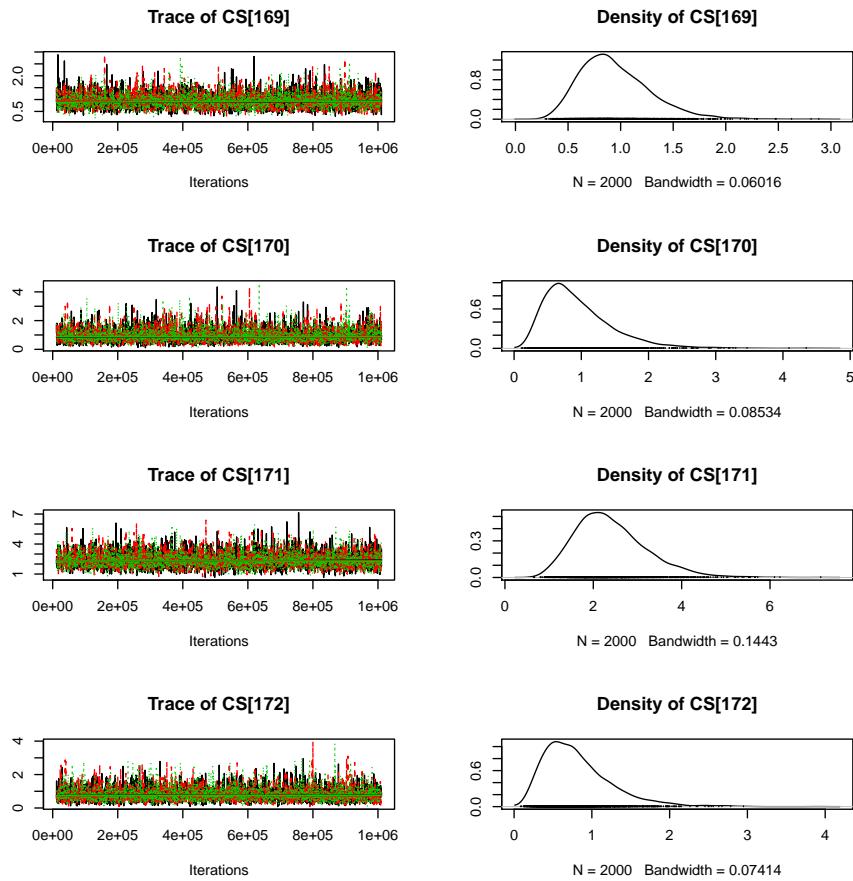


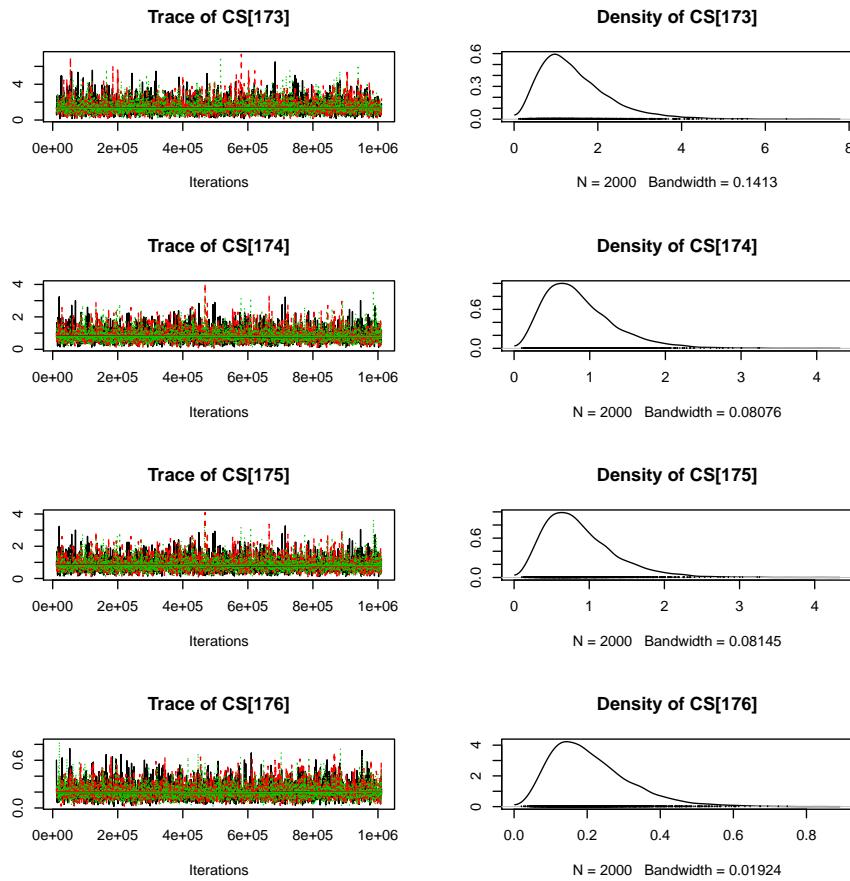


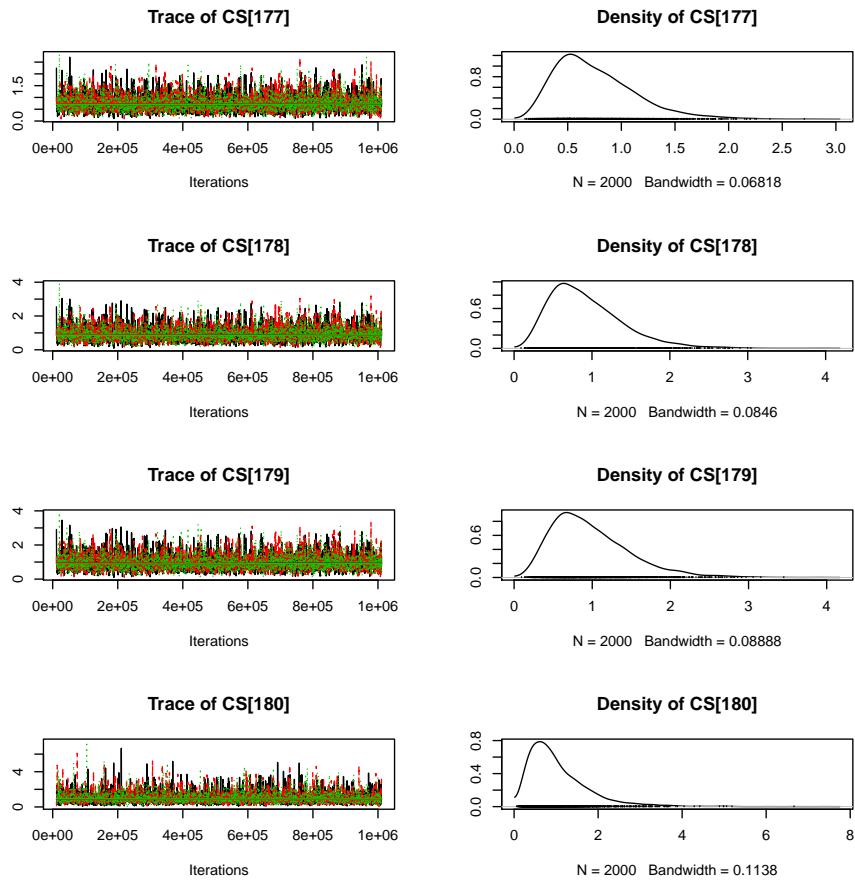


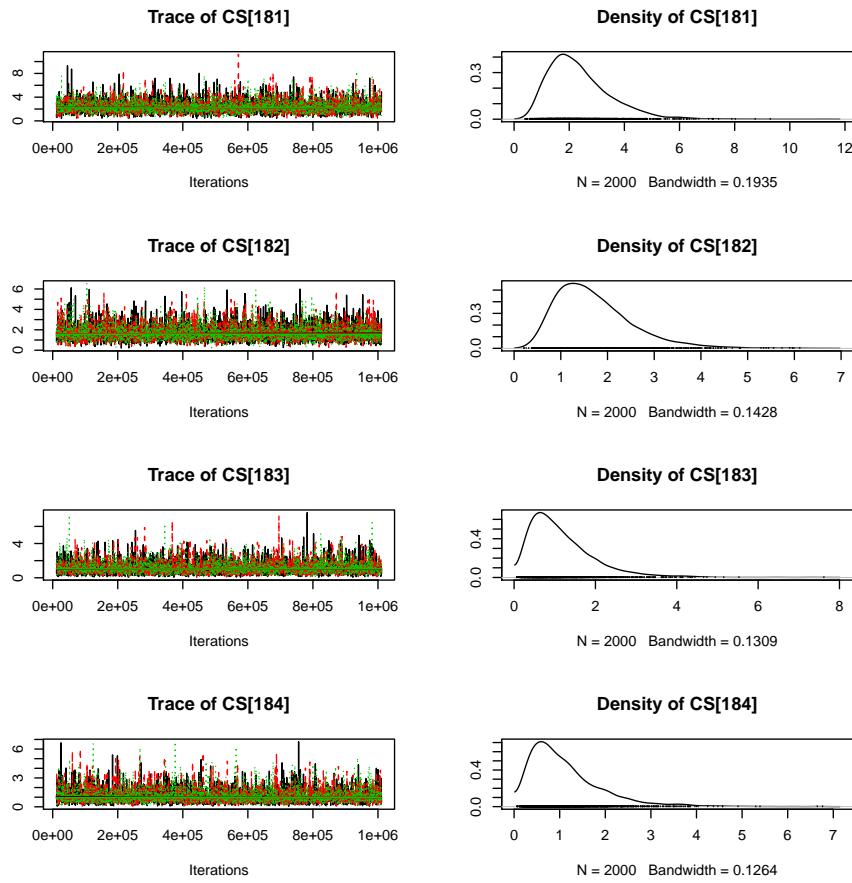


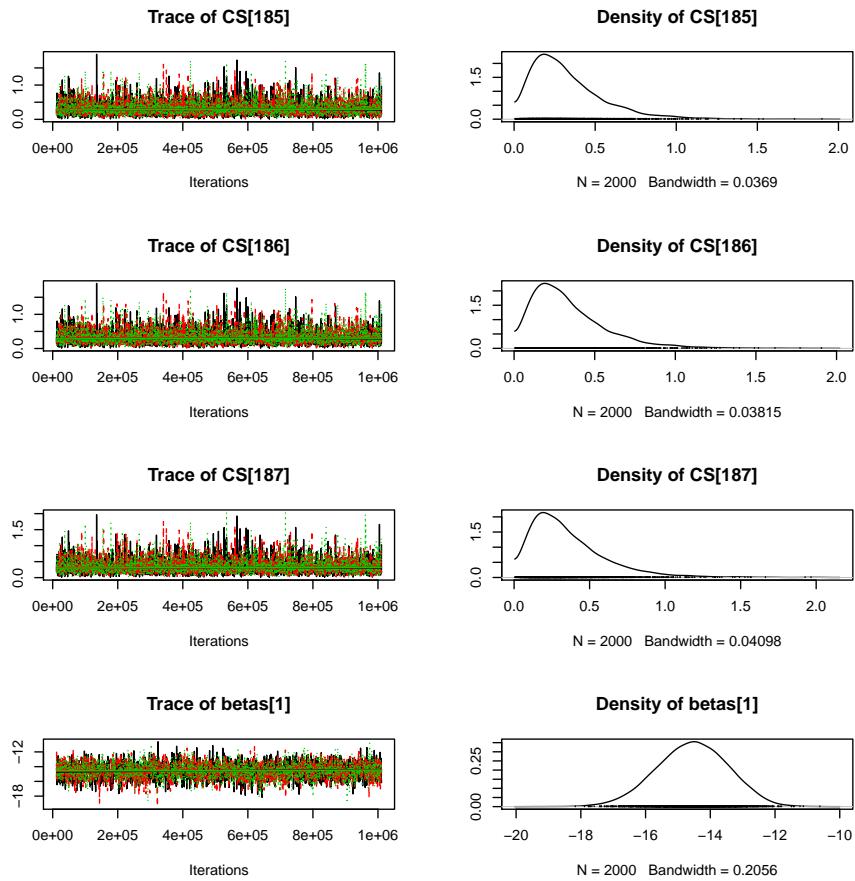


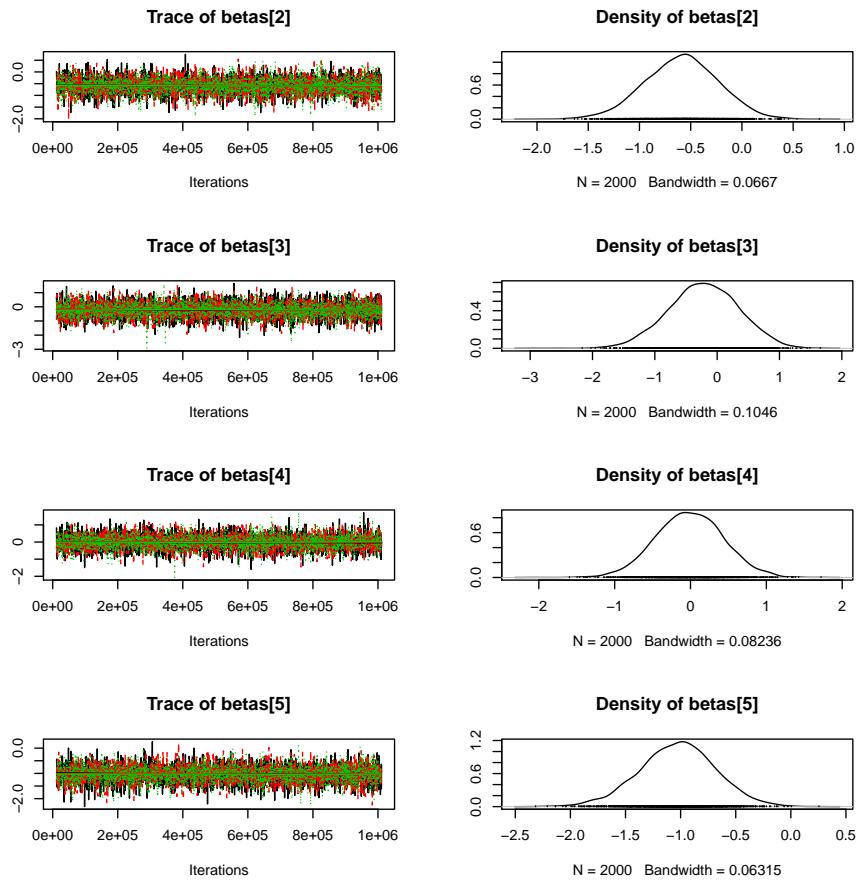


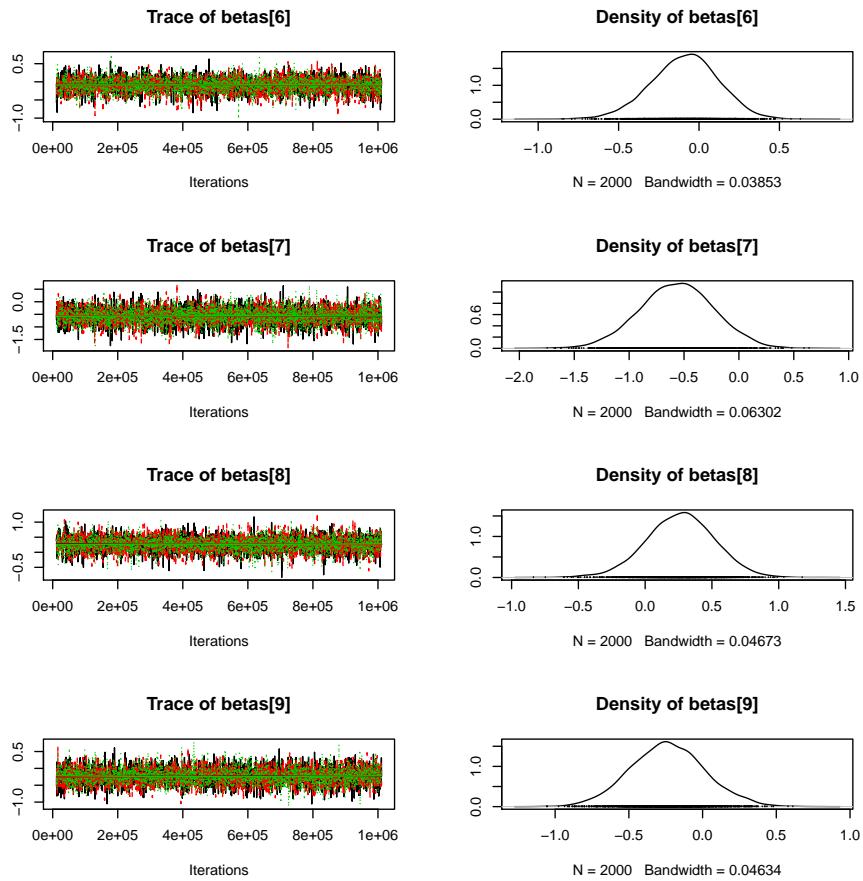


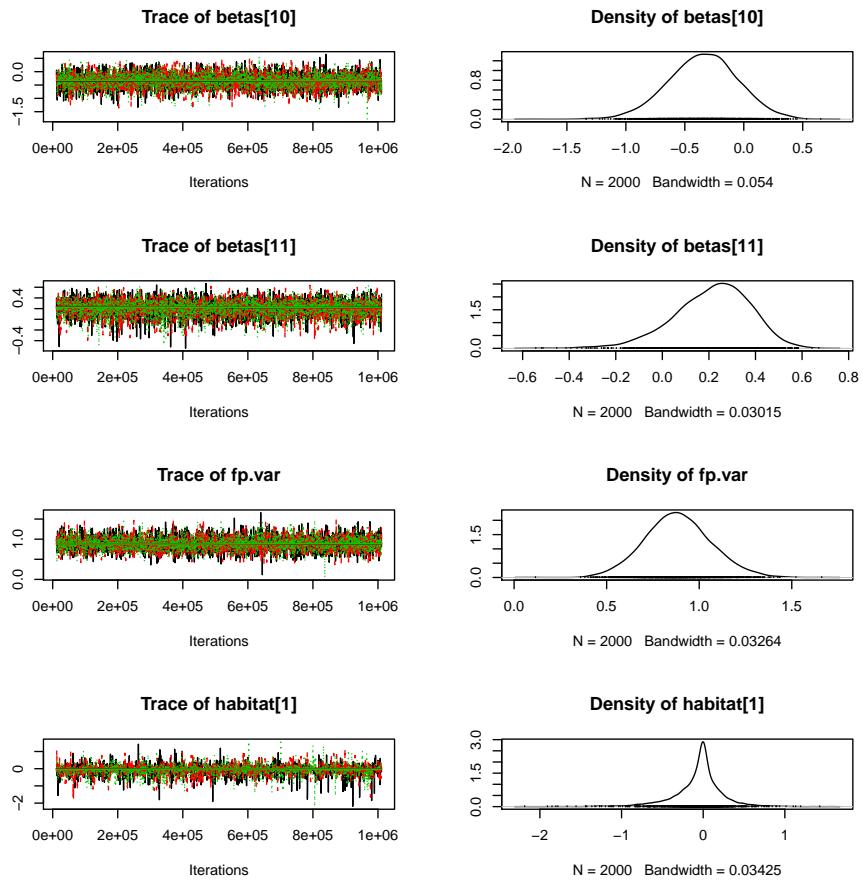


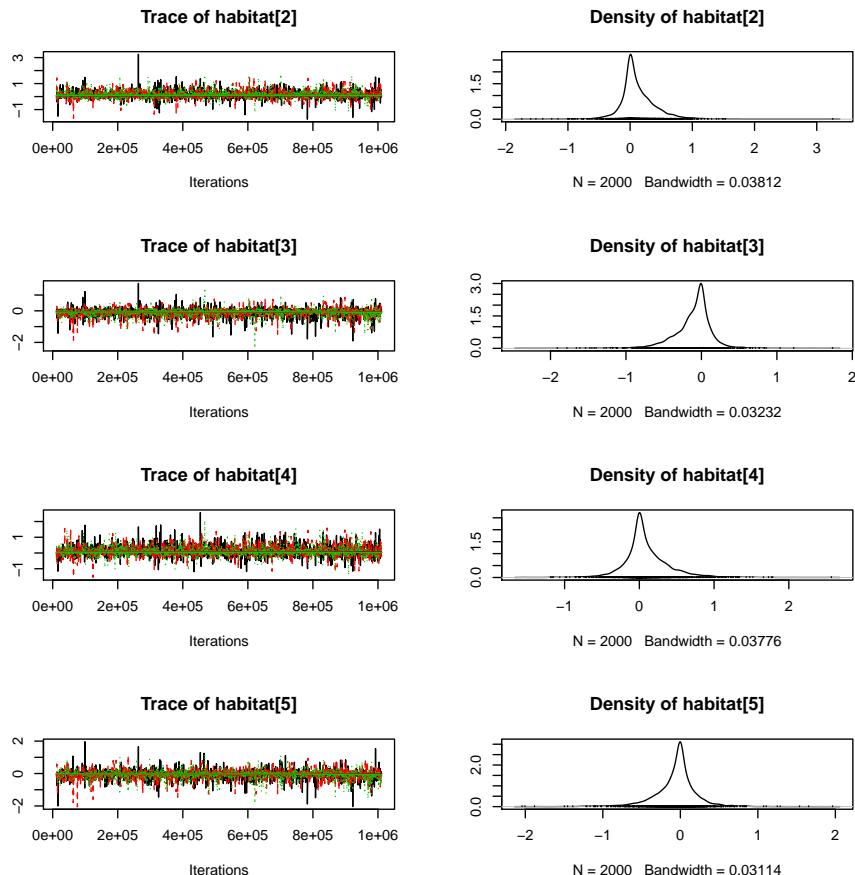




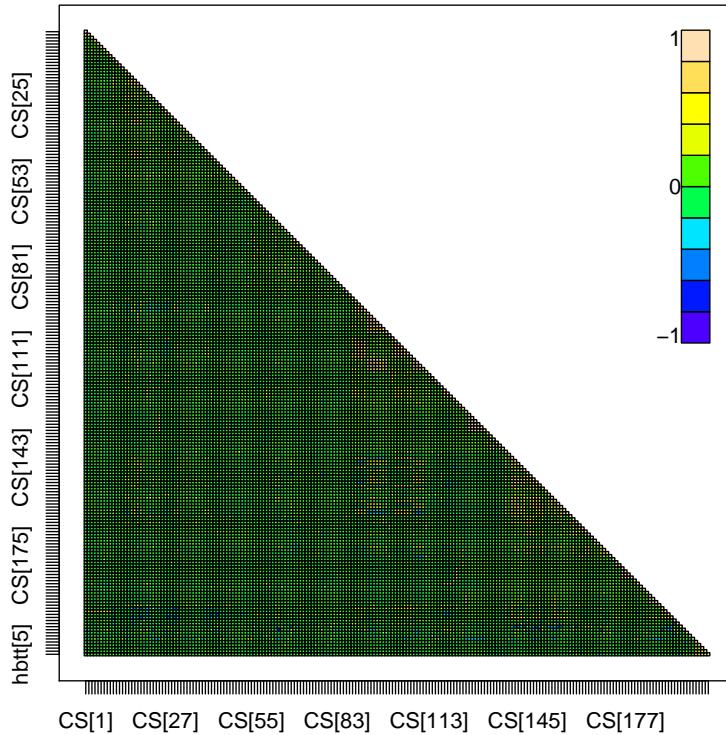








```
crosscorr.plot(a.out)
```



```

# get coeffs from the chains - pull in some helper functions and Rdata from
# Bayesian model
source("helper_functs.R")
# load('~/Work/Dropbox/First year of assessment/FA_V001.RData')

# get posterior for cox-snell(CS) residuals from MCMC
CS.full <- tbl_df(get_coef_chains(model.out = a.out, coef.names = "CS"))

# just look at mean CS for now, can put posterior around it later
CS.means <- CS.full %>% group_by(Parameter) %>% summarise(post.mean = mean(MCMC))

library(survival)

## Loading required package: splines

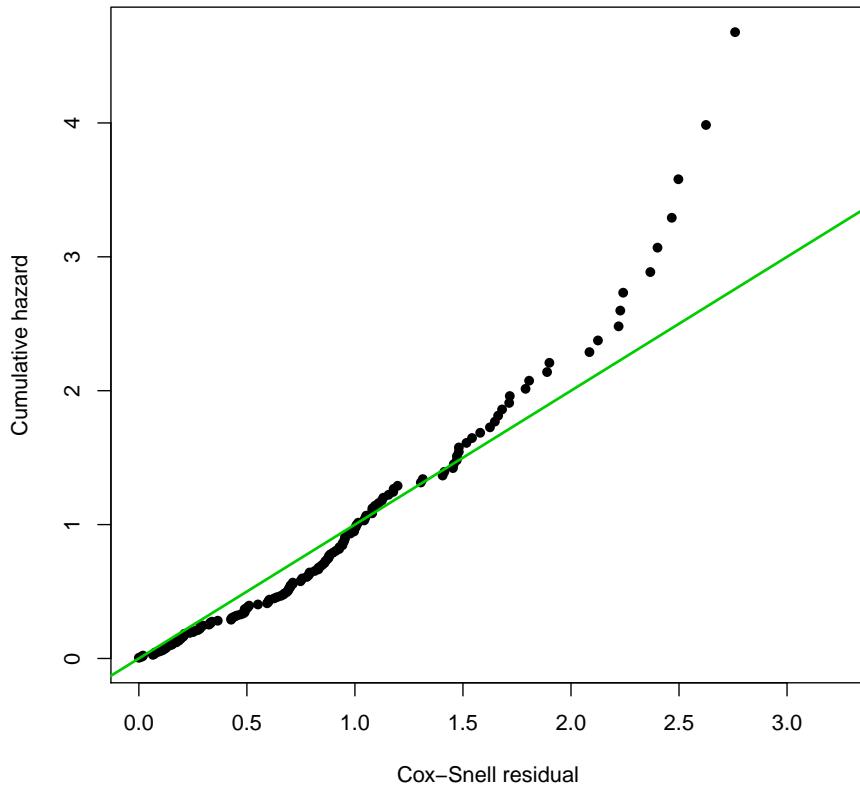
CS = CS.means$post.mean

```

```

# Kaplan-Meyer non-parametric survival at CS - should follow exp(1)
# distribution
km.cs <- survfit(Surv(CS, !censored) ~ 1)
summary.km.cs <- summary(km.cs)
rcu <- summary.km.cs$time # Cox-Snell residuals of
# uncensored points.
surv.cs <- summary.km.cs$surv
plot(rcu, -log(surv.cs), type = "p", pch = 16, xlab = "Cox-Snell residual",
     ylab = "Cumulative hazard")
abline(a = 0, b = 1, col = 3, lwd = 2)

```



It looks as though the fit of the Weibull isn't too bad, some deviation is expected in the tails of the distribution, but over the bulk it seems to follow the 1:1 line fairly closely. Perhaps not much of an improvement over the normal GLM above - should try different distributional assumptions here. We could also plot the posterior for the CS residuals and see if the 1:1 line falls outside a

Table 1: Posterior mean and  $P(\beta > 0)$  for model parameters

Parameter	Posterior Mean	Bayesian P
landings	-0.21	0.10
Lmax..cm.	0.09	0.66
price	0.35	0.88
Recreational.pc.catch	-0.26	0.15
RegionAlaska	0.59	0.95
RegionUSEC	0.22	0.65
RegionUSSE	0.01	0.51
RegionUSWC	1.03	1.00
TL	0.57	0.95
X.Intercept.	14.56	1.00
Year.of.fishery.development..stock.based.	0.24	0.83

Table 2: Posterior mean and  $P(\beta > 0)$  for model habitat

Habitat	Posterior Mean	Bayesian P
benthic	0.06	0.57
benthopelagic	-0.12	0.33
demersal	0.11	0.67
pelagic	-0.09	0.39
reef	0.06	0.58

95% interval for any points.

```
coeffs <- tbl_df(get_coef_chains(model.out = a.out, coef.names = "betas", var.names = colnames(a.out))

# regressin coeffs are -beta
coef_P <- coeffs %>% group_by(Parameter) %>% summarise(post.mean = -mean(MCMC),
  post.P = 1 - mean(MCMC > 0))

# habitat
habs <- tbl_df(get_coef_chains(model.out = a.out, coef.names = "habitat", var.names = with(y,
  levels(factor(habitat_MM)))))

hab_P <- habs %>% group_by(Parameter) %>% summarise(post.mean = -mean(MCMC),
  post.P = 1 - mean(MCMC > 0))

# finite population variance of family random effects
fp.vars <- tbl_df(get_coef_chains(model.out = a.out, coef.names = "fp.var")) %>%
  summarise(mean(MCMC)^2)

fp.vars
```

```
## Source: local data frame [1 x 1]
##
##   mean(MCMC)^2
## 1      0.7839
```