

# Calibration

AUTHOR

Tom Gibbens-Matsuyama

For this assignment, I decided to use a high flow metric.

```
library(tidyverse)
library(here)

msage = readRDS(here("data/msage.RDS"))
#View(msage)

# first rearrange so we can plot all results
msage1 = msage %>% gather(key="run",value="str", -date, -month, -day, -year, -wy,-obs)
source(here("R/compute_highflowmetrics.R"))

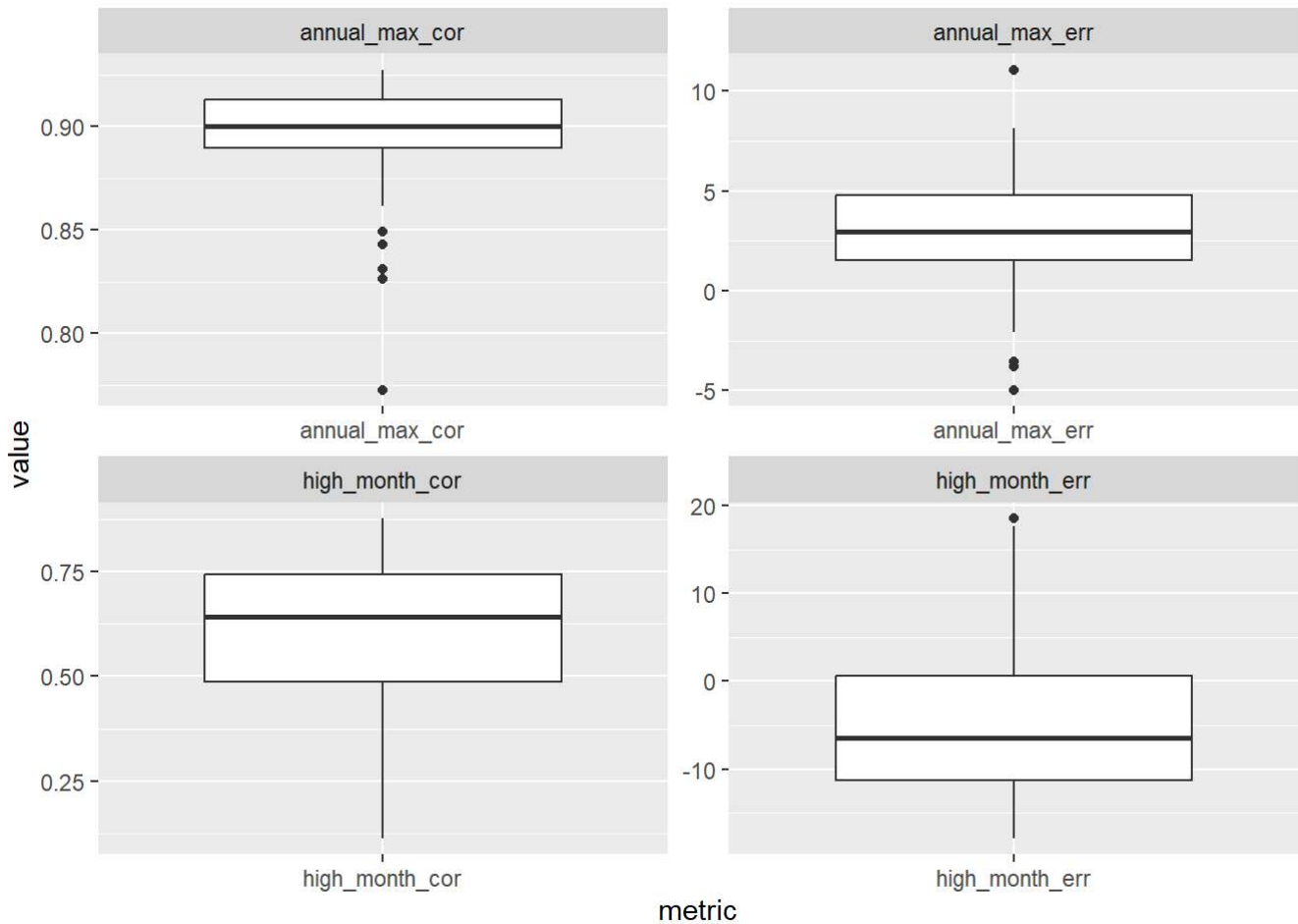
# another example using our low flow statistics
# use apply to compute for all the data
res = msage %>% select(-date, -month, -day, -year, -wy, -obs ) %>% apply(2,FUN=compute_highflowme

# extract information from the list
results = as.data.frame(matrix(unlist(res), byrow=T, ncol=4))
colnames(results)=c("annual_max_err","annual_max_cor", "high_month_err",
                    "high_month_cor")

# interesting to look at range of metrics - could use this to decide on
# acceptable values
summary(results)
```

annual_max_err	annual_max_cor	high_month_err	high_month_cor
Min. : -4.985	Min. : 0.7723	Min. : -17.9036	Min. : 0.1124
1st Qu.: 1.524	1st Qu.: 0.8898	1st Qu.: -11.2773	1st Qu.: 0.4870
Median : 2.941	Median : 0.9003	Median : -6.4274	Median : 0.6434
Mean : 2.974	Mean : 0.8964	Mean : -5.1683	Mean : 0.5949
3rd Qu.: 4.796	3rd Qu.: 0.9130	3rd Qu.: 0.5664	3rd Qu.: 0.7439
Max. : 11.066	Max. : 0.9274	Max. : 18.5729	Max. : 0.8783

```
# graph range of performance measures
resultsl = results %>% gather(key="metric",value="value")
ggplot(resultsl, aes(metric, value))+geom_boxplot()+facet_wrap(~metric, scales="free")
```



## Parameter Set

```
# pick parameter set with greater low flow month correlation
```

```
best_par2 = which.max(results$high_month_cor)
```

```
# first we need to be able to identify parameter sets in results
```

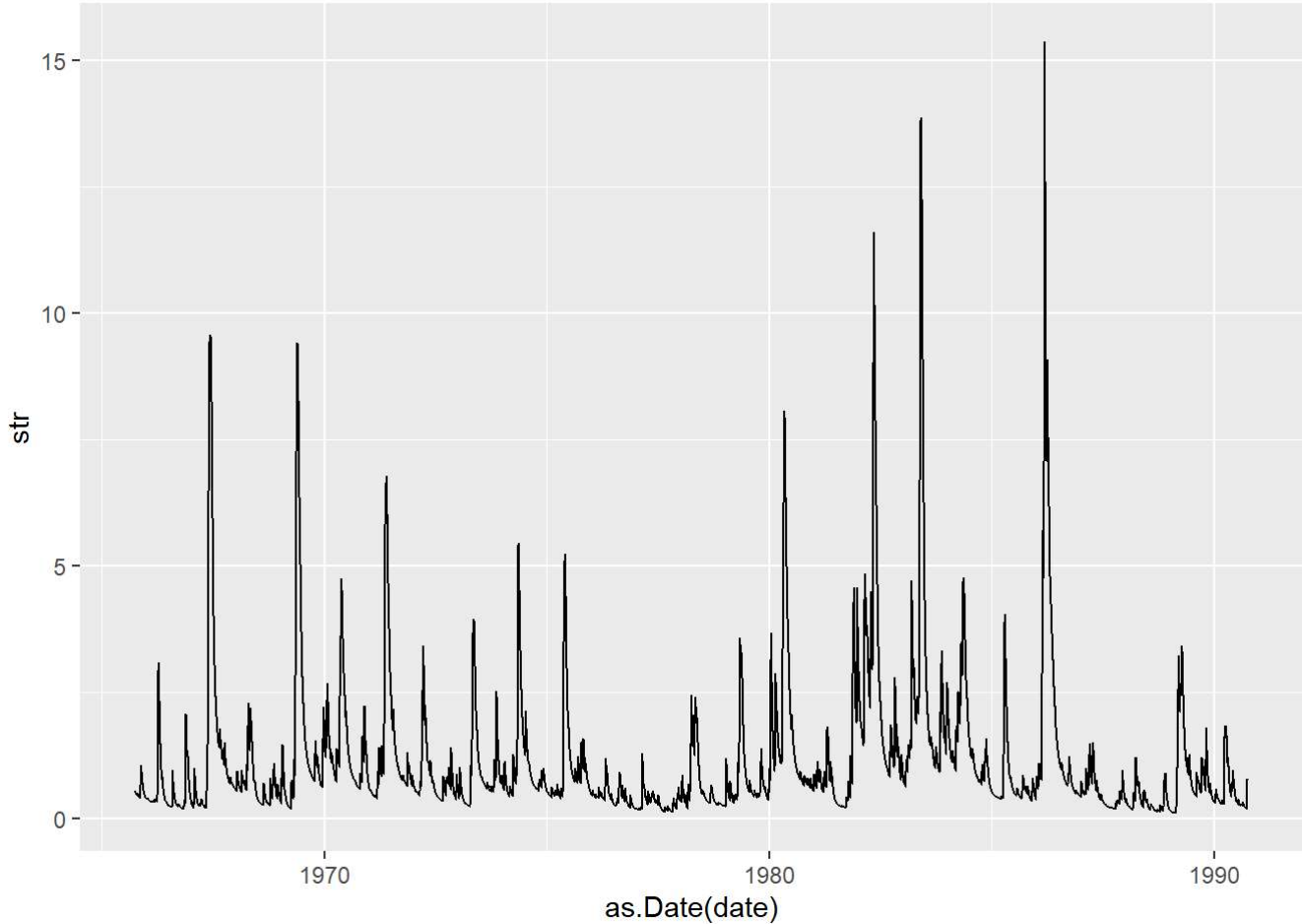
```
ID = msage %>% select(-date, -month, -day, -year, -wy, -obs ) %>% colnames()
results$id = ID
```

```
# now we can pick only good parameters
```

```
accept_par = subset(results, annual_max_cor > 0.7 & high_month_cor > 0.7 & abs(annual_max_err < 0
nrow(accept_par)
```

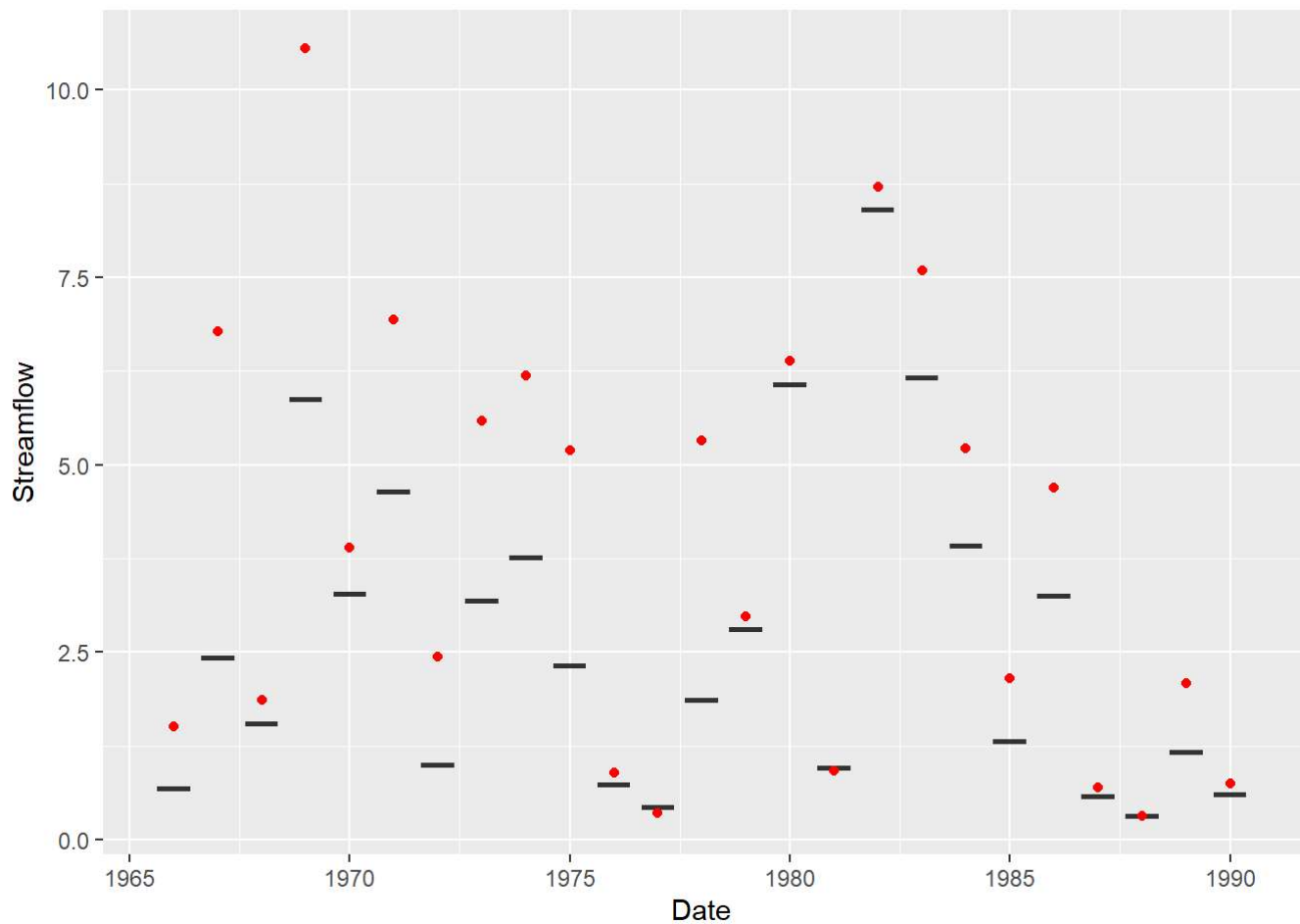
```
[1] 1
```

```
# plot these to compare with all parameters
msagel$accept = ifelse(msagel$run %in% accept_par$id, TRUE, FALSE)
ggplot(subset(msagel, accept), aes(as.Date(date), str))+geom_line()
```



```
# or with observed but focus on August flow
msagel_mth = msagel %>% select(-date,-day) %>% group_by(month, wy, run, accept) %>% summarise(str)

# just august flow
ggplot(subset(msagel_mth, month==5 & accept),
  aes(wy, str, group=wy))+geom_boxplot(position="dodge")+
  geom_point(aes(wy, obs), col="red")+labs(y="Streamflow", x="Date")
```



```
mean_jan_obs = mean(subset(msagel_mth, month==5)$obs)
ggplot(subset(msagel_mth, month==5), aes(accept, str))+geom_boxplot()+
  geom_hline(yintercept=mean_jan_obs, col="red")+labs(y="Streamflow", x="Date")
```

