QAQC of the CINP ASSP 1994 - 2018 CPUE data

EKelsey

4/10/2020

Load libraries

```
library(dplyr)

## Warning: package 'dplyr' was built under R version 3.6.1

##
## 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

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.6.1
```

Load data and add columns neccessary for QAQC:

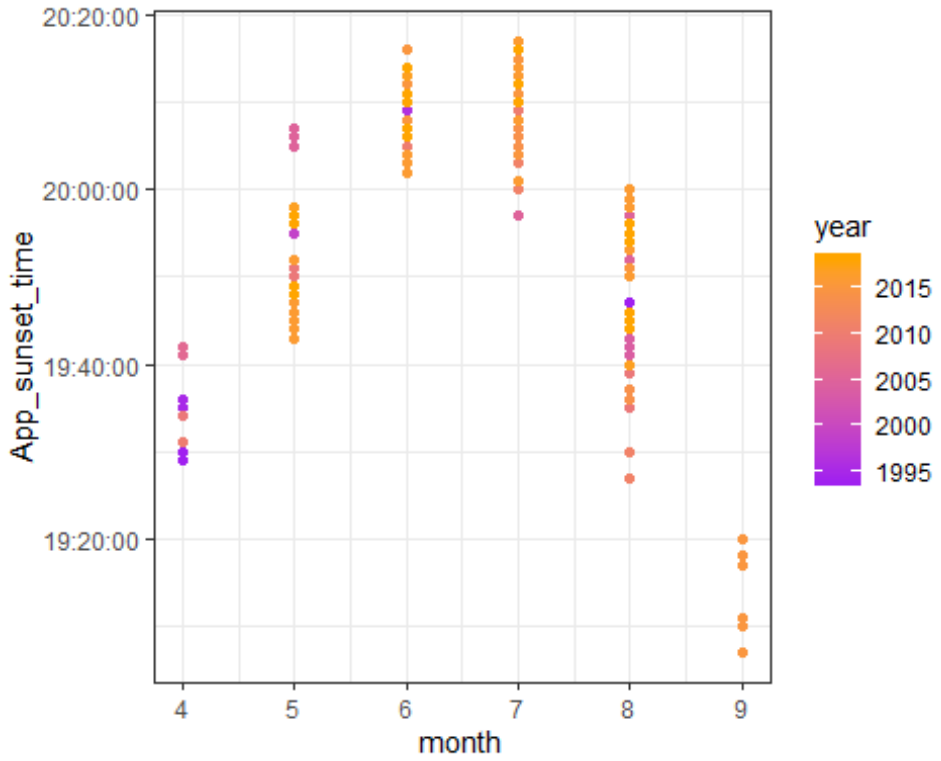
```
metadata <- read.csv('~/WERC-
SC/ASSP_share/ASSP_4_metadata_CPUE_20200325.csv') %>%
  mutate_at(c("App_sunset", "std_ending"), .funs = ~as.POSIXct(.,
format="%m/%d/%Y %H:%M")) %>%
  mutate_at(c("net_open_1", "net_close_1", "net_open_2", "net_close_2",
"net_open_3",
              "net_close_3", "net_open_4", "net_close_4", "net_open_5",
"net_close_5"),
            .funs = ~as.POSIXct(., format="%Y-%m-%d %H:%M:%S")) %>%
  mutate_at(c("App_sunset", "std_ending", "net_open_1", "net_close_1"),
            .funs = list(time = ~ hms::as_hms(.))) %>%
  mutate(CPUE_ratio = CPUEstd/CPUEraw) %>%
  filter(TRUE)
```

Graphical check of App_sunset

Plotted by month

```
ggplot(metadata, aes(month, App_sunset_time)) +
  geom_point(aes(color = year)) +
```

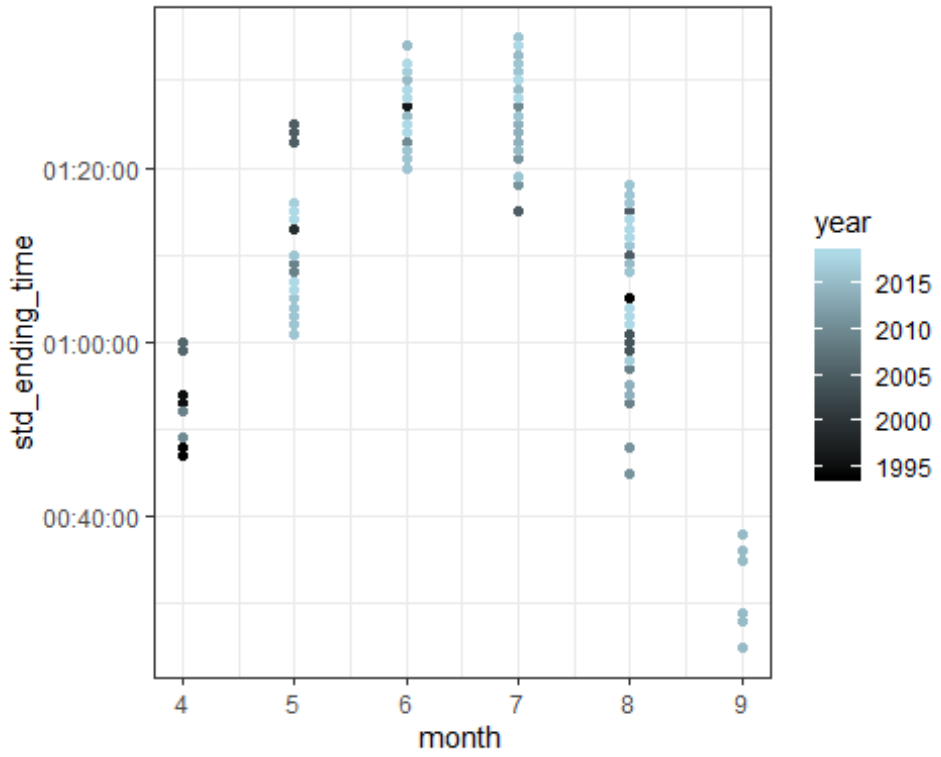
```
  scale_color_gradient(low="purple", high="orange") +
  theme_bw()
```



Graphical check of Std_ending

Plotted by month

```
ggplot(metadata, aes(month, std_ending_time)) +
  geom_point(aes(color = year)) +
  scale_color_gradient(low="black", high="light blue") +
  theme_bw()
```



Summarize net_open and net_close (not perfect because don't account for next-day effort)

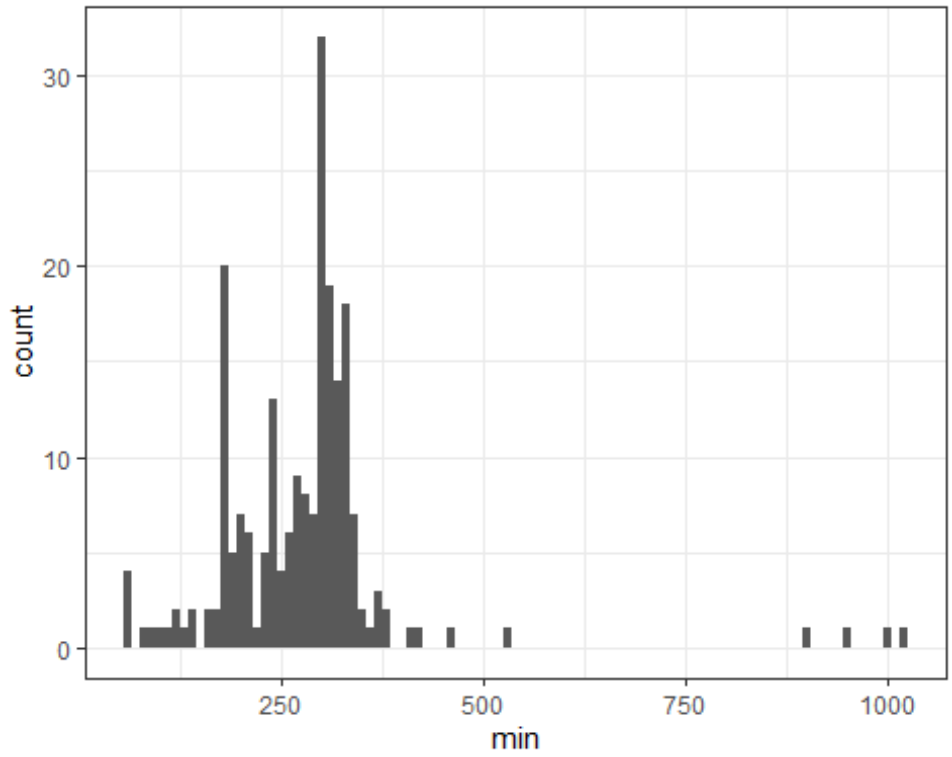
```
# first (and usually only) net open time
summary(as.POSIXct(metadata$net_open_1_time))

##                   Min.                 1st Qu.                  Median
## "1970-01-01 00:00:00" "1970-01-01 20:45:00" "1970-01-01 21:02:30"
##                   Mean                 3rd Qu.                    Max.
## "1970-01-01 20:33:23" "1970-01-01 21:36:00" "1970-01-01 23:35:00"
##                   NA's
##                   "22"

# first (and usually only) net close time
summary(as.POSIXct(metadata$net_close_1_time))

##                   Min.                 1st Qu.                  Median
## "1970-01-01 00:00:00" "1970-01-01 01:24:00" "1970-01-01 02:00:00"
##                   Mean                 3rd Qu.                    Max.
## "1970-01-01 03:37:47" "1970-01-01 02:17:30" "1970-01-01 23:59:00"
##                   NA's
##                   "21"
```

—

Total mistnetting minutes per session

```
library(ggplot2)
ggplot(metadata, aes(min)) +
  geom_histogram(binwidth = 10) +
  theme_bw()
```

```
## Warning: Removed 22 rows containing non-finite values (stat_bin).
```




```
# summary of total mistnetting minutes
summary(metadata$min)
```

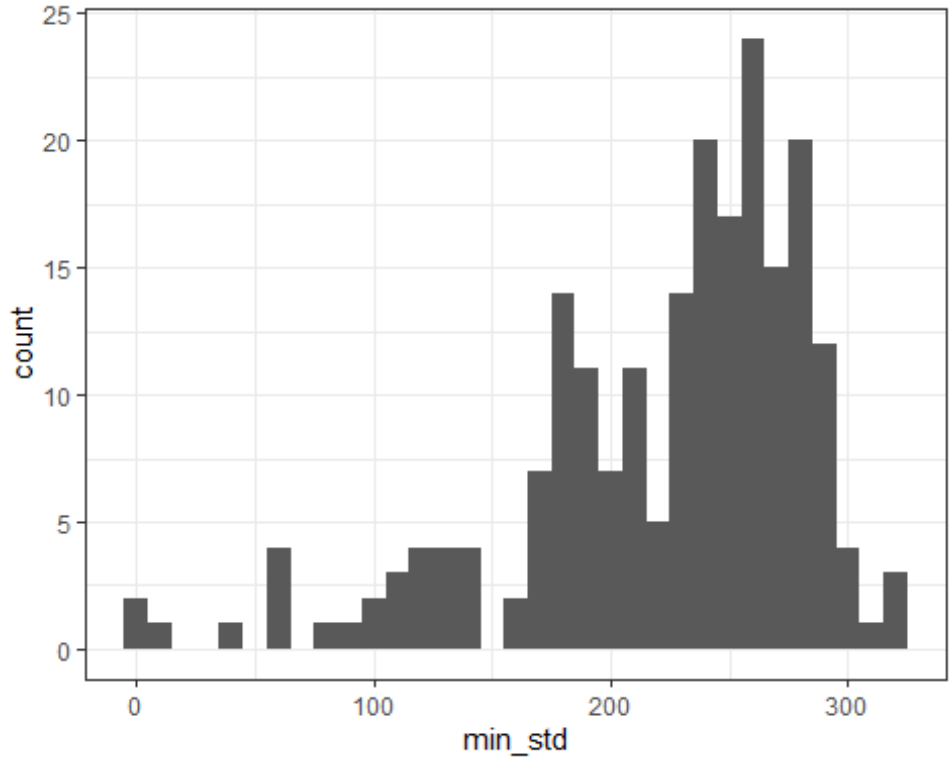
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    56.0   214.0   293.0   280.3   316.0  1022.0      22
```

Total mistnetting standard minutes per session

from start until end or standard ending, whichever came first

```
ggplot(metadata, aes(min_std)) +
  geom_histogram(binwidth = 10) +
  theme_bw()
```

```
## Warning: Removed 22 rows containing non-finite values (stat_bin).
```



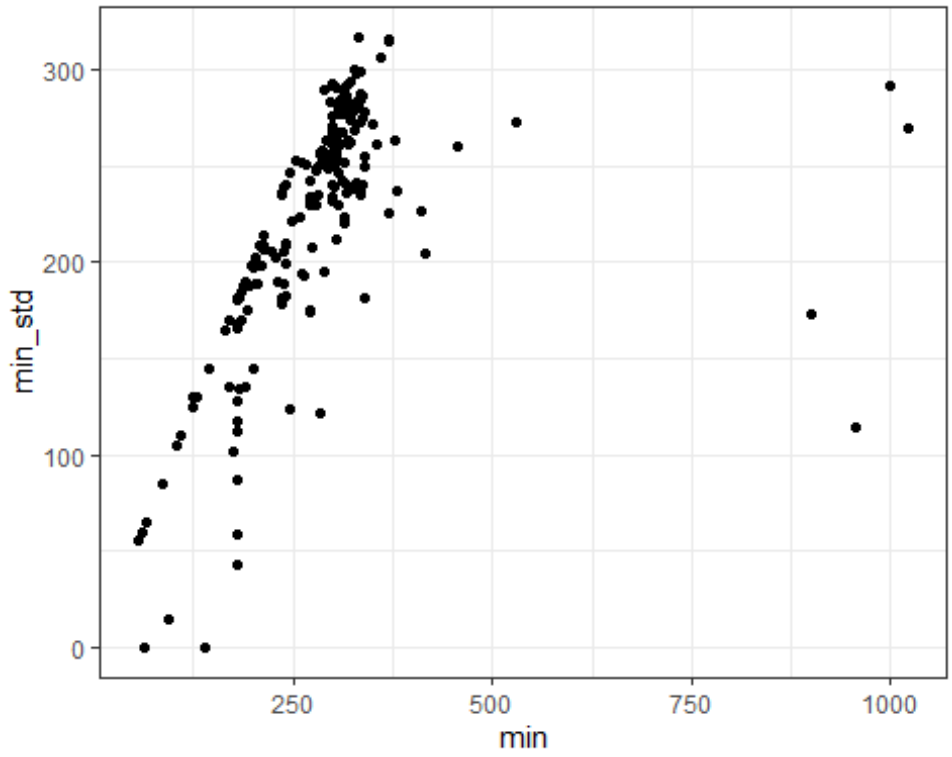

```
# summary of mistnetting minutes cut to standard ending time
summary(metadata$min_std)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##     0.0   189.0   239.6   223.0   267.3   317.1      22
```

Compare min vs. min_std for each session

```
library(ggplot2)
ggplot(metadata, aes(min, min_std)) +
  geom_point() +
  theme_bw()
```

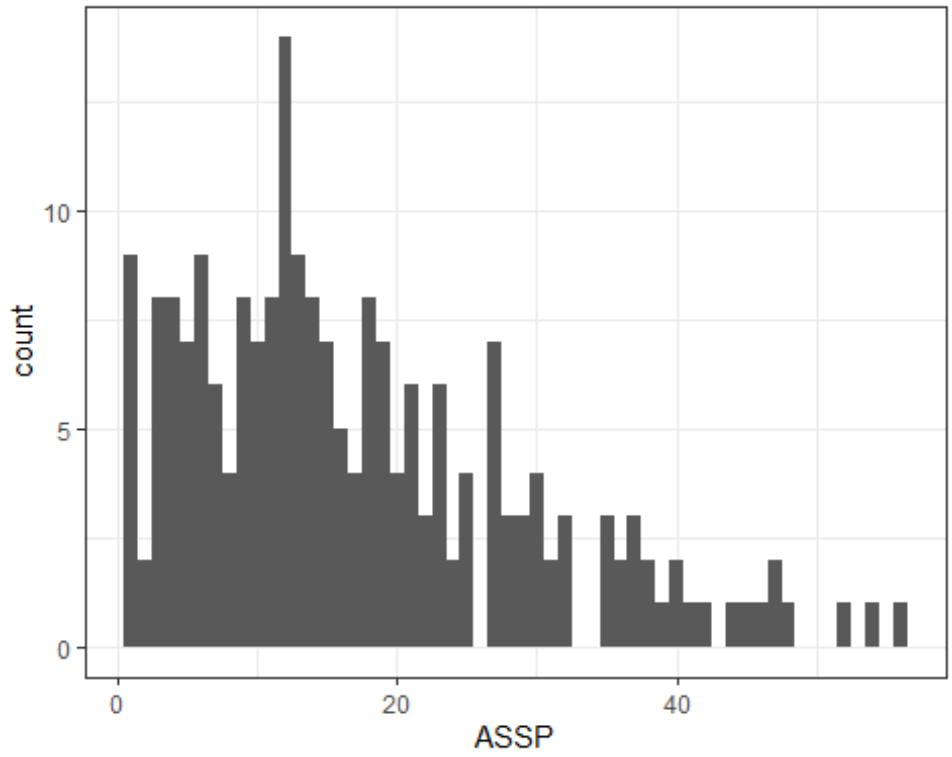
```
## Warning: Removed 22 rows containing missing values (geom_point).
```



—

Histogram of total ASSP caught per session

```
ggplot(metadata, aes(ASSP)) +
  geom_histogram(binwidth = 1) +
  theme_bw()

## Warning: Removed 27 rows containing non-finite values (stat_bin).
```



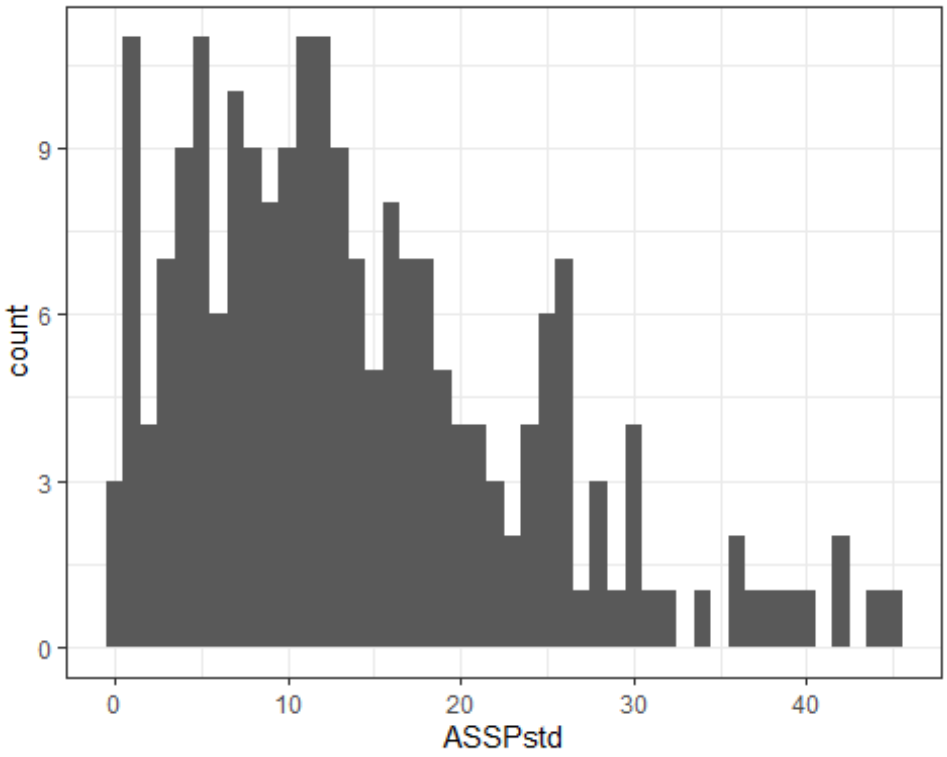

```
# summary of ASSP catches
summary(metadata$ASSP)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    1.00    8.00   14.00   17.18   23.00   56.00      27
```

Histogram of total ASSP caught per standardized session

```
ggplot(metadata, aes(ASSPstd)) +
  geom_histogram(binwidth = 1) +
  theme_bw()
```

```
## Warning: Removed 27 rows containing non-finite values (stat_bin).
```



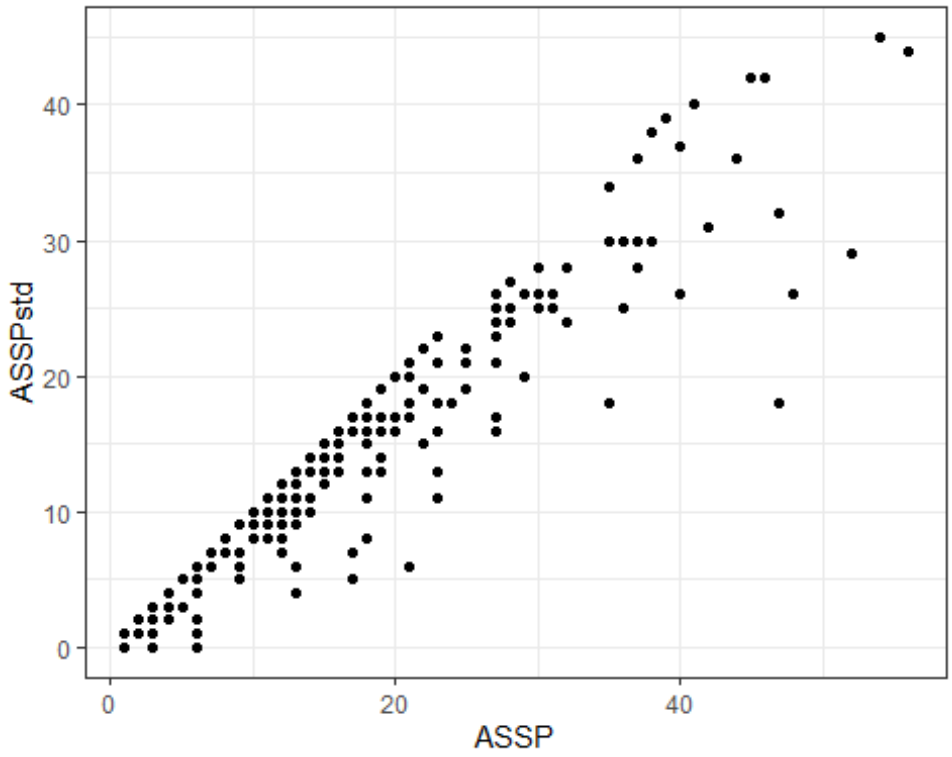

```
# summary of standardized ASSP catches
summary(metadata$ASSPstd)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    0.00    7.00   12.00   14.08   19.00   45.00      27
```

comparison of ASSP vs ASSPstd

```
ggplot(metadata, aes(ASSP, ASSPstd)) +
  geom_point() +
  theme_bw()
```

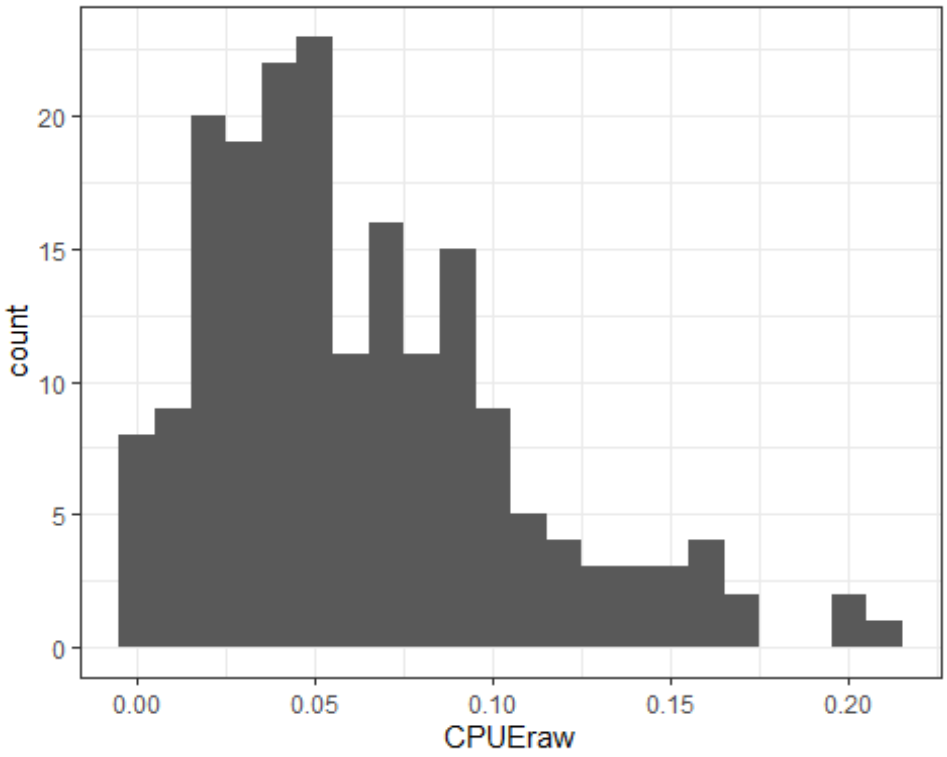
```
## Warning: Removed 27 rows containing missing values (geom_point).
```



—

visualization of CPUE per session

```
ggplot(metadata, aes(CPUEraw)) +
  geom_histogram(binwidth = 0.01) +
  theme_bw()

## Warning: Removed 46 rows containing non-finite values (stat_bin).
```



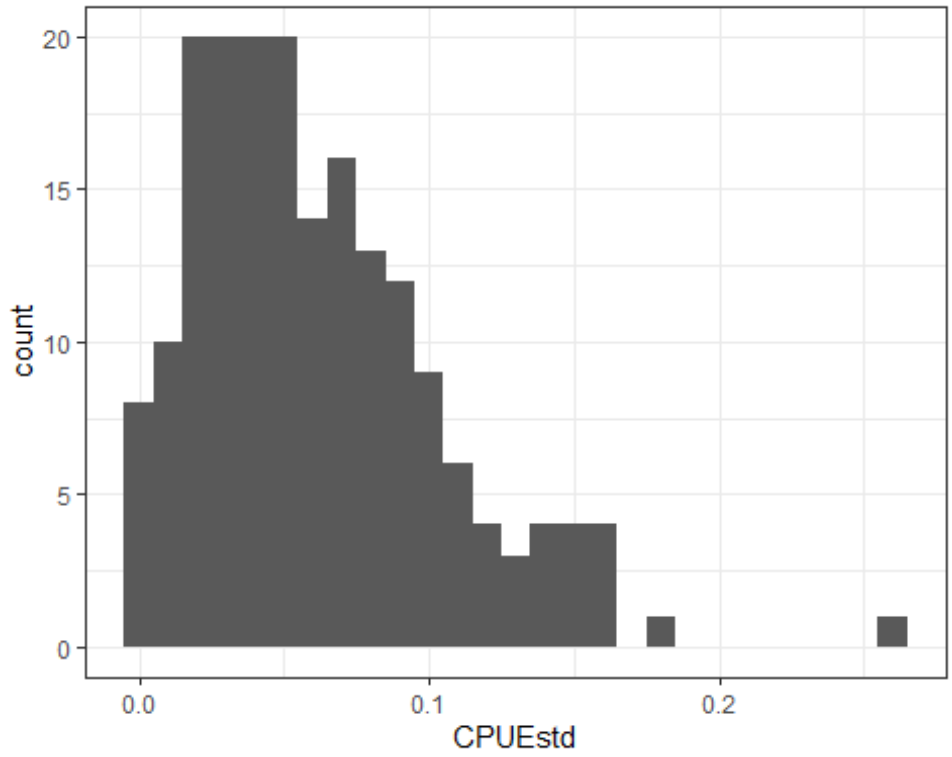

```
# summary of CPUE per session
summary(metadata$CPUEraw)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## 0.00282 0.03245 0.05298 0.06270 0.08795 0.20779      46
```

visualization of CPUE per standardized session

```
ggplot(metadata, aes(CPUEstd)) +
  geom_histogram(binwidth = 0.01) +
  theme_bw()
```

```
## Warning: Removed 47 rows containing non-finite values (stat_bin).
```



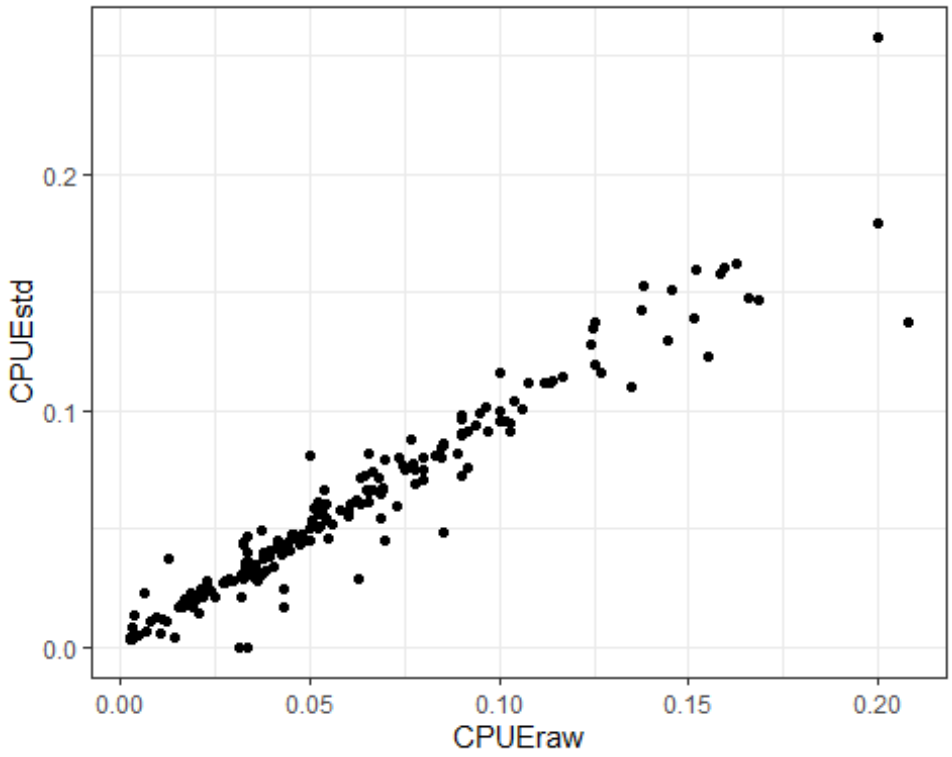

```
# summary of CPUE per standardized session
summary(metadata$CPUEstd)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## 0.00000 0.02858 0.05351 0.06166 0.08511 0.25810      47
```

comparision of CPUE vs CPUEstd

```
ggplot(metadata, aes(CPUEraw, CPUEstd)) +
  geom_point() +
  theme_bw()
```

```
## Warning: Removed 47 rows containing missing values (geom_point).
```


CPUEstd
0.2
0.1
0.0
0.00
0.05
0.10
0.15
0.20
CPUEraw