EMP Data 2024

Catarina Pien

2025-02-10

Goal

Examine USBR stations listed in the Water Quality Control Plan to determine whether we are meeting requirements.

Notes

Communications from Ted

- The station contact is Melanie Lowe at USBR's California-Great Basin Regional Office
- When we last emailed in December, she said "our office staff clean and calibrate the stations monthly when we are able to access them, and Central Valley Office coordinates with Tracy Office regarding issues w/access or stations going down/not reporting. However, I believe Tracy Office is currently cleaning/calibrating/maintaining the Collinsville and Ripon stations routinely."
- C19 discontinued due to safety concerns and because the city of Vallejo no longer uses the intake on Cache Slough. Ted was not able to find any documentation that this was ever officially communicated to SWRCB.

Requirements in Table 5 of the WQCP

- 15-minute Electrical Conductivity and Water Temperature at CLL, SAL, CNT, UNI, PCT, CCS, EMM
- 1 to 15-minute EC, WT, pH, DO, turbidity, chlorophyll a fluorescence, tidal elevation and meteorological data (air temperature, wind speed and direction, solar radiation) for DMC

What is on CDEC

- C2 – Sacramento River at Collinsville (CLL): EC (100 - event), WT (25 - hourly)

- C4 – SJR at San Andreas Landing (SAL): EC (100 - event), WT (25 - hourly)
- C5 – Contra Costa Canal at Pumping Plant #1 (CNT): EC (100 - event), WT (25 - hourly)
- C8 – Old River near Middle River (UNI): EC (100 - event), WT (25 - hourly)
- C14 – Sacramento River at Port Chicago (PCT): EC (sensor 100), WT (25 - event)
- *C19 – Cache Slough at City of Vallejo Intake (CCS): discontinued 1/1/2015
- D22 – Sacramento River at Emmaton (EMM): EC (100 - event), WT (25 - hourly)
- DMC1 – Delta-Mendota Canal at Jones Pumping Plant (DMC): EC (100 - event), WT (25 - hourly)

Conclusions

- We are not quite collecting at the frequency specified by Table 5 for water temperature.
- We are not collecting all the parameters specifed by Table 5 for DMC.
- Most stations seem to be collected relatively well for 2024, except CNT which was missing a few months of data.
- CLL and DMC also have >10% data outside of suitable range/not collected.

Next Steps

- We could get the QA/QC protocols from Melanie's group. All we did here was basic range filtering, and not even 100% sure of appropriate EC range.
- We could check if extra data are collected that are not reported on CDEC.
- We could check why certain parameters are collected at frequencies different from what has been specified in the Water Quality Control Plan

Analysis

Electrical Conductivity

Download data

Examine range of values

```r
ggplot(ec %>% filter(!is.na(month))) +
  geom_boxplot(aes(month,value, fill = month))+
  facet_wrap(~station_id, scales = "free_y", ncol = 2) +
  labs(y = "EC (uS/cm)")+
  theme_bw() +
  theme(legend.position = "bottom")
```
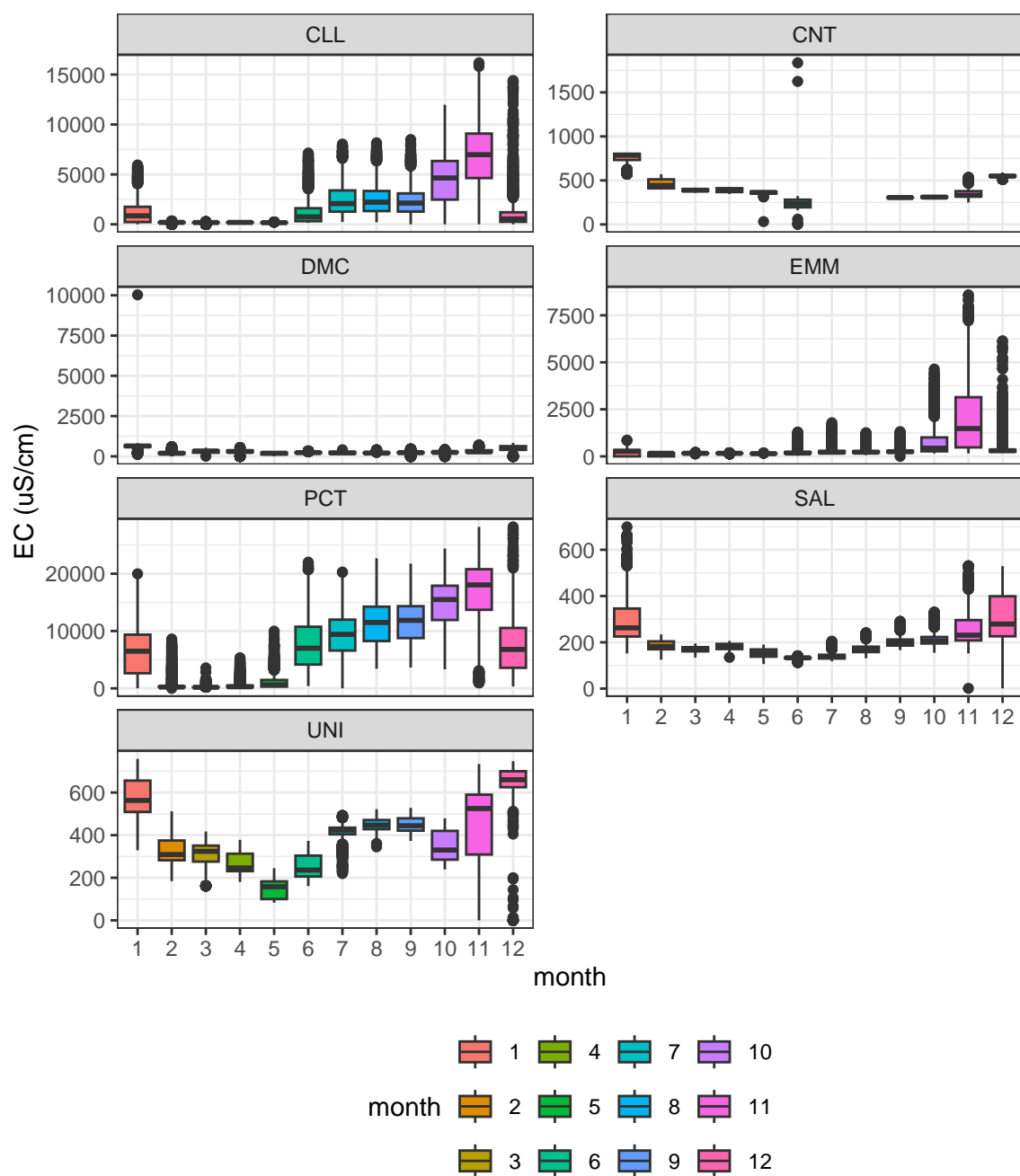
Figure 1: Boxplot of EC (uS/cm) values in 2024.

Filtered to suitable ranges, with 1000 uS/cm as upper limit for CNT, DMC, SAL, UNI, and
no upper limit for other stations.

Histogram for filtered values

```
ggplot(ec %>%filter(range == 1)) + geom_histogram(aes(value), color = "lightblue4")+
  facet_wrap(~station_id, scales = "free_x", ncol = 2) +
  labs(x = "EC (uS/cm)")+
  theme_bw()
```
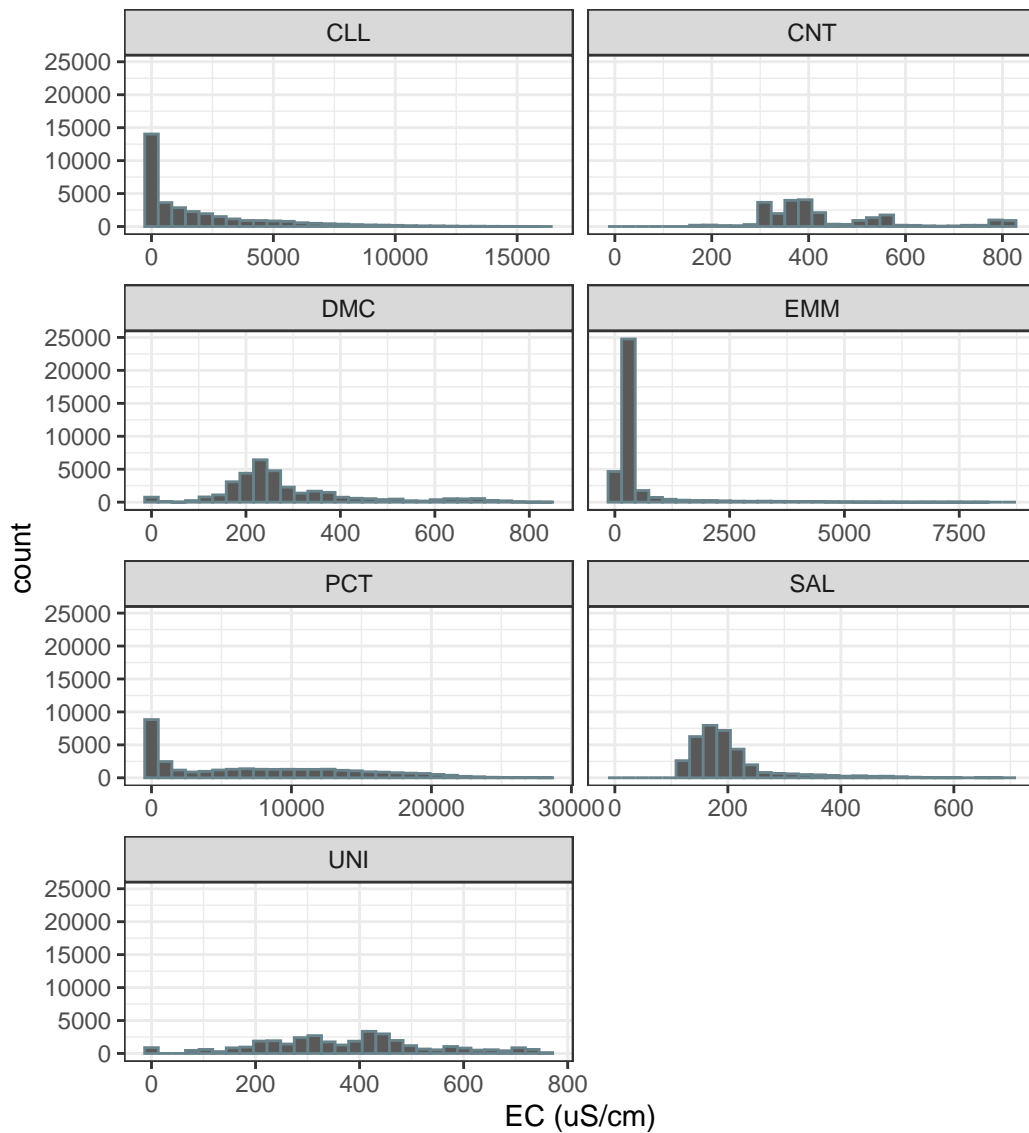

Figure 2: Histogram of EC values.

Water Temperature

Download Data

Examine range of water temperature values

```r
ggplot(wt %>% filter(!is.na(month))) +
  geom_boxplot(aes(month, value, fill = month))+
  facet_wrap(~station_id, scales = "free_y", ncol = 2)+
  labs(y = "Water Temperature (F)")+
  theme_bw()+
  theme(legend.position = "bottom")
```
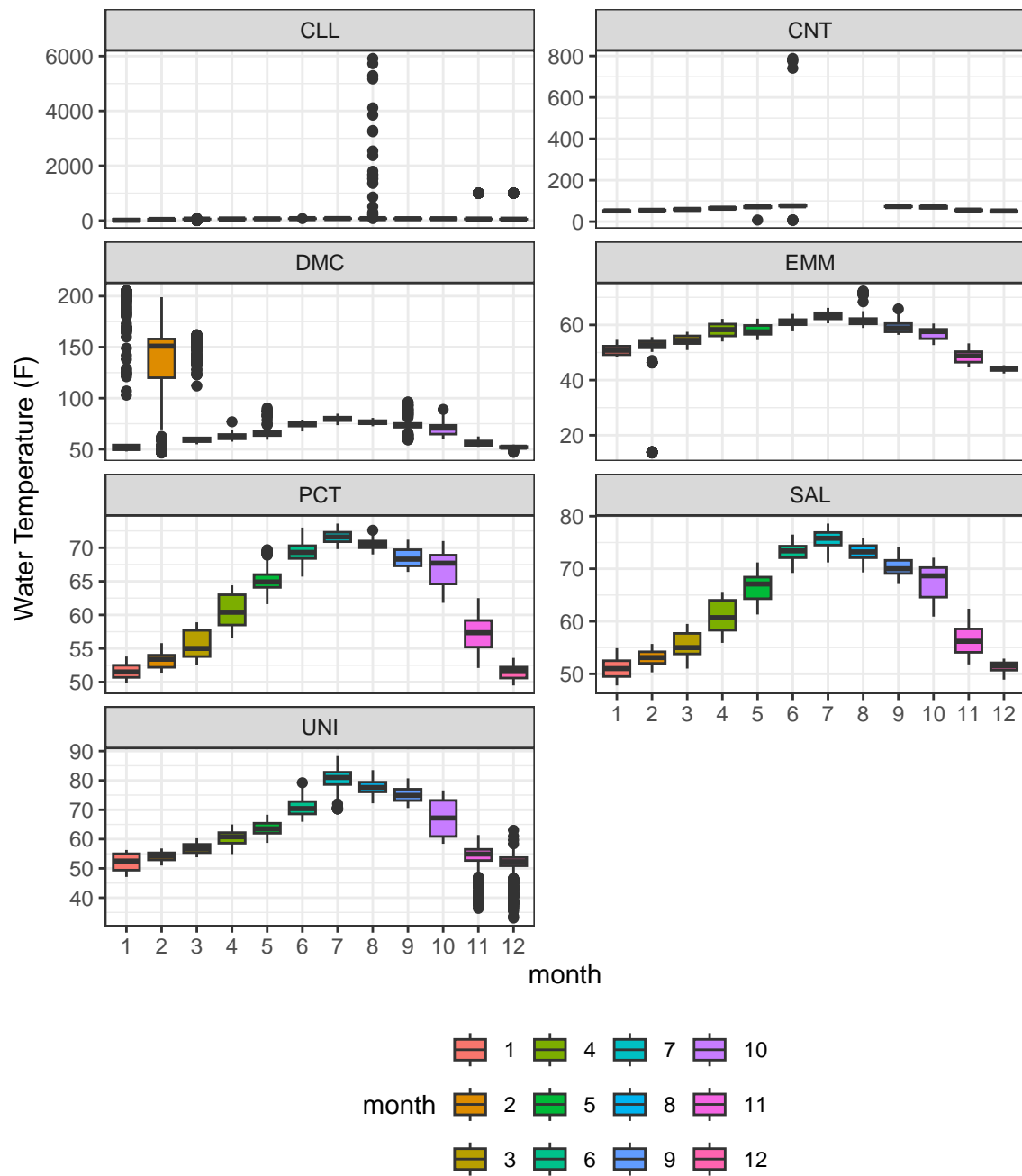
Figure 3: Boxplot of WT (deg F) values in 2024.

Filtered to suitable ranges, with 30F as lower limit and 100F as upper limit. Not sure what the deal with DMC is in February, CLL in August.

Filtered data histogram

```
ggplot(wt %>%filter(range == 1)) + geom_histogram(aes(value), color = "lightblue4")+
  facet_wrap(~station_id, ncol = 2) +
  labs(x = "Water Temperature (F)")+
  theme_bw()
```
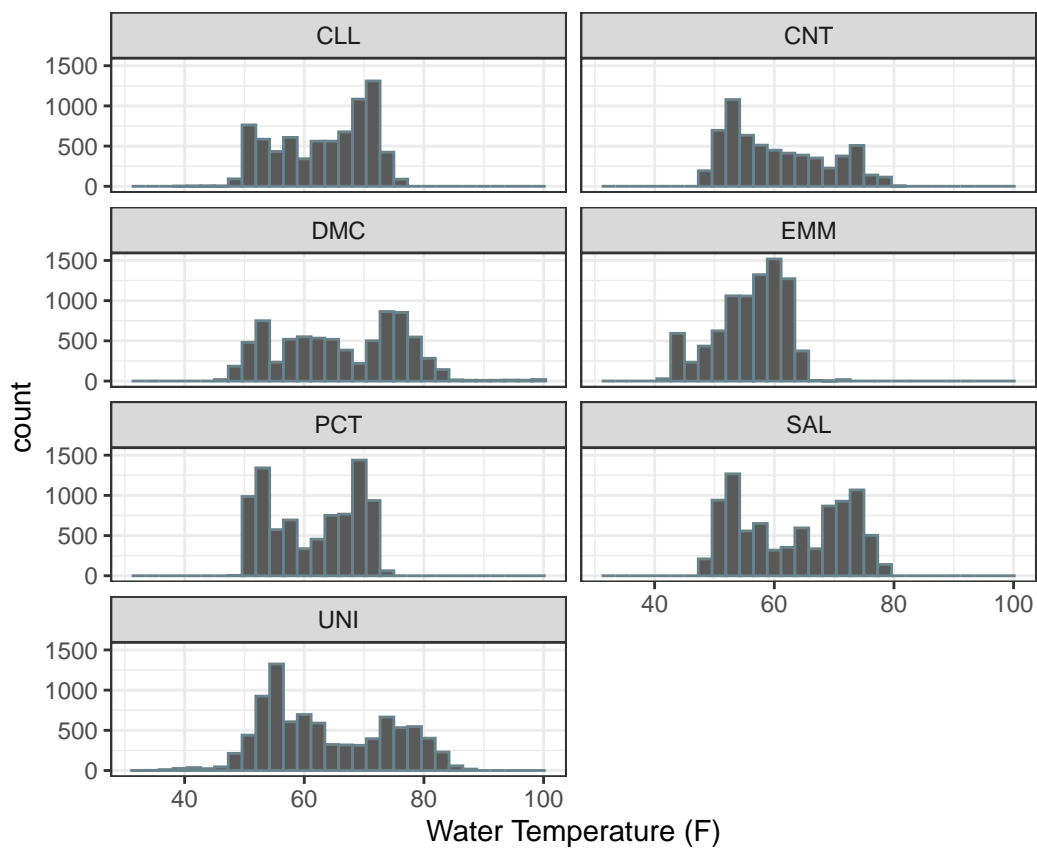

Figure 4: Histogram of WT values.

Annual Summary

```r
prop_ec_year <- ec %>%
  group_by(station_id) %>%
  summarize(n = n(),
            max = max(value),
            min=min(value),
            sum_good = sum(range, na.rm = TRUE),
            prop_good = round(sum_good/n,3), # data within range/data collected
            prop_all = round(sum_good/(24*4*366),3))%>%#data within range/data sho
  mutate(sensor = "EC")
```

```r
prop_wt_year <- wt %>%
  group_by(station_id) %>%
  summarize(n = n(),
            max = max(value),
            min=min(value),
            sum_good = sum(range, na.rm = TRUE),
            prop_good = round(sum_good/n,3),
            prop_all = round(sum_good/(24*366),3)) %>%
  mutate(sensor = "Temperature")
```

```r
prop_data <- bind_rows(prop_ec_year, prop_wt_year)
```

Summary

```r
ggplot(prop_data) +
  geom_col(aes(station_id, prop_all, fill = sensor), position = position_dodge2()) +
  scale_fill_manual(values = c("orange2", "steelblue4")) +
  labs(x = "Station Code", y = "Proportion Data Within Reasonable Range") +
  scale_y_continuous(breaks = seq(0, 1, by = 0.1)) +
  theme_bw()
```
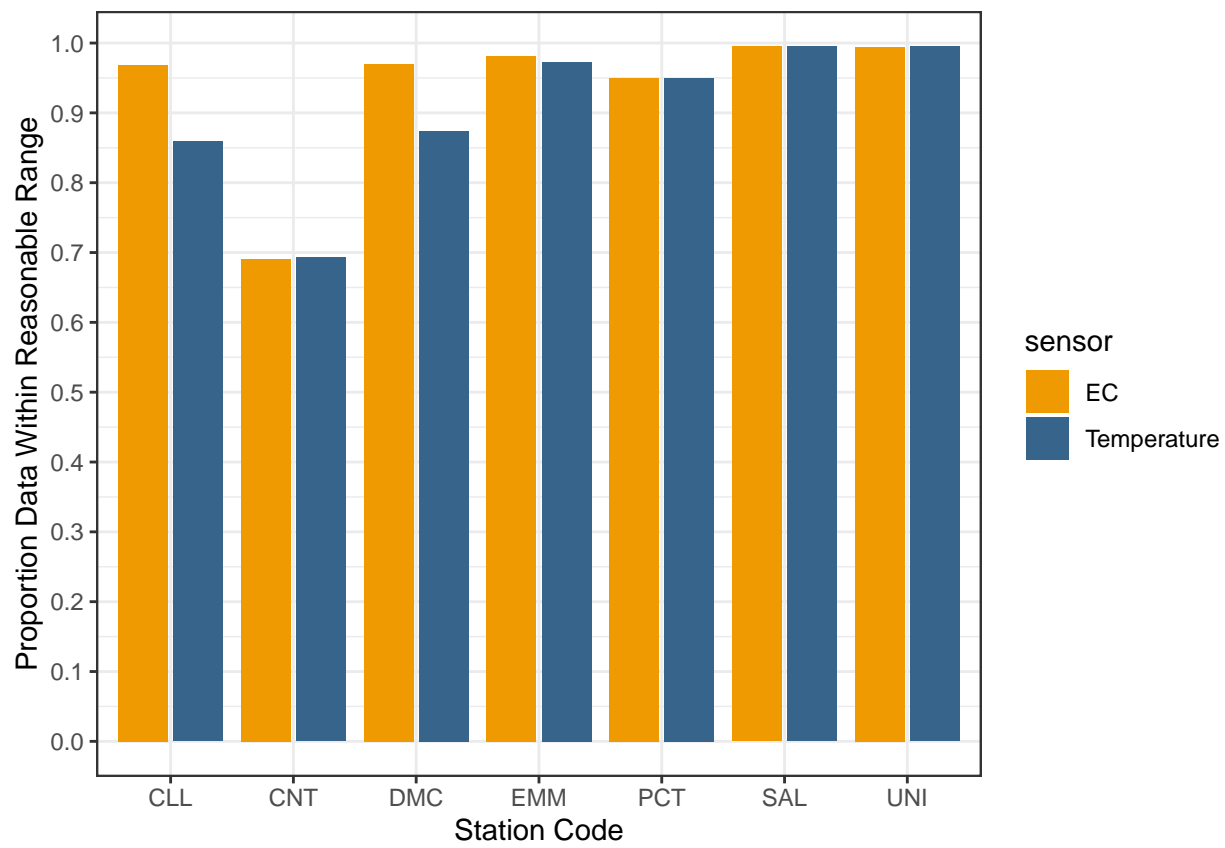
Figure 5: Proportion of EC and WT Data Collected and within Range in 2024.

```
library(kableExtra)
kableExtra::kable(prop_data)
```

Table 1: Sample Sizes of Data and Proportion of Good
Data. Prop_good = Proportion of data that was col-
lected that was within range. Prop_all = Proportion of
data that should have been collected (if all time points
were available) that was within range.

station_id	n	max	min	sum_good	prop_good	prop_all	sensor
CLL	34044	16177.0	0.0	34044	1.000	0.969	EC
CNT	24287	1836.0	1.0	24285	1.000	0.691	EC
DMC	34067	10028.0	0.0	34066	1.000	0.970	EC
EMM	34463	8588.0	0.0	34463	1.000	0.981	EC
PCT	33388	28175.0	0.0	33388	1.000	0.950	EC
SAL	34947	699.0	1.0	34947	1.000	0.995	EC
UNI	34937	758.0	0.0	34937	1.000	0.994	EC
CLL	8510	5914.0	8.5	7546	0.887	0.859	Temperature
CNT	6097	788.0	6.1	6087	0.998	0.693	Temperature
DMC	8520	205.0	46.4	7679	0.901	0.874	Temperature
EMM	8618	72.3	13.6	8548	0.992	0.973	Temperature
PCT	8348	73.6	49.5	8348	1.000	0.950	Temperature
SAL	8738	78.6	47.8	8738	1.000	0.995	Temperature
UNI	8739	88.3	33.1	8739	1.000	0.995	Temperature

Individual Station Plots

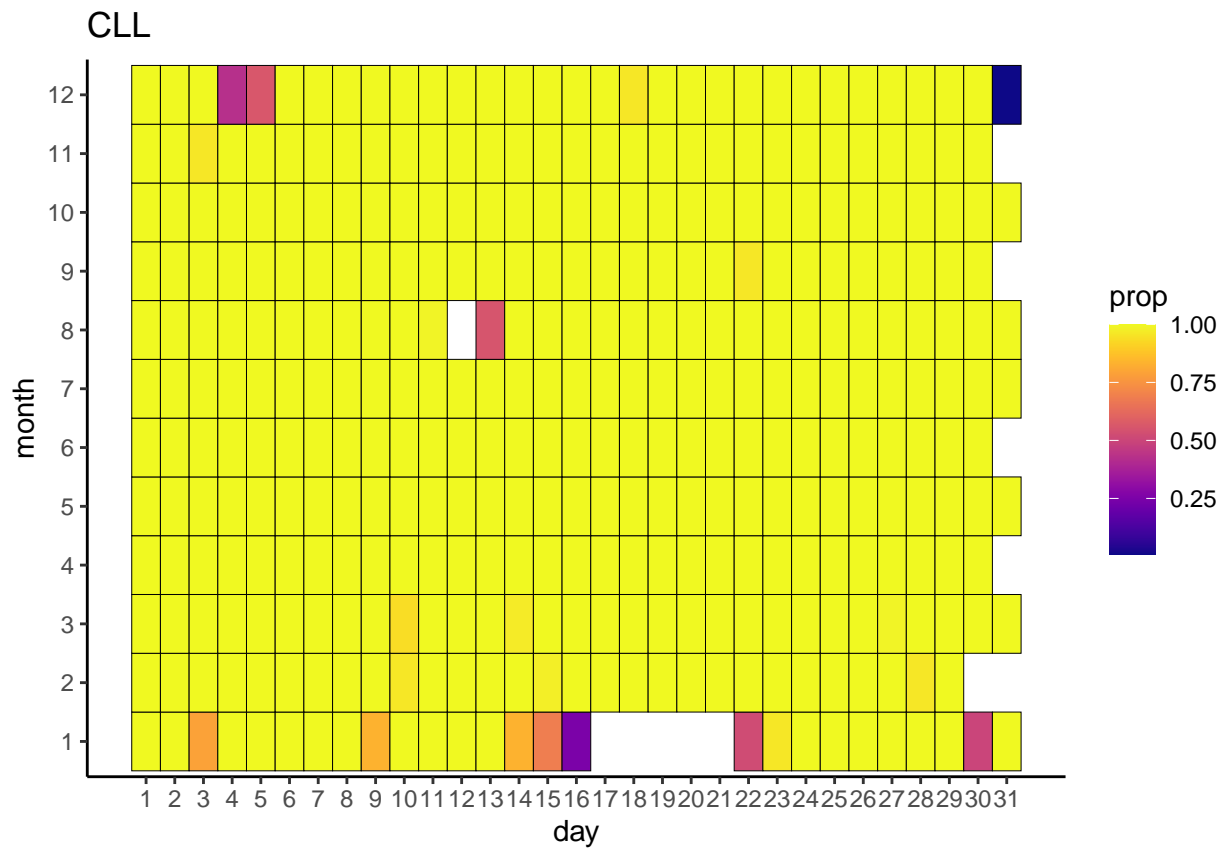## Electrical Conductivity

```
plot_ec("CLL")
```


Figure 6: Proportion of EC values collected and within range by day at CLL. White boxes are missing data for the entire day.
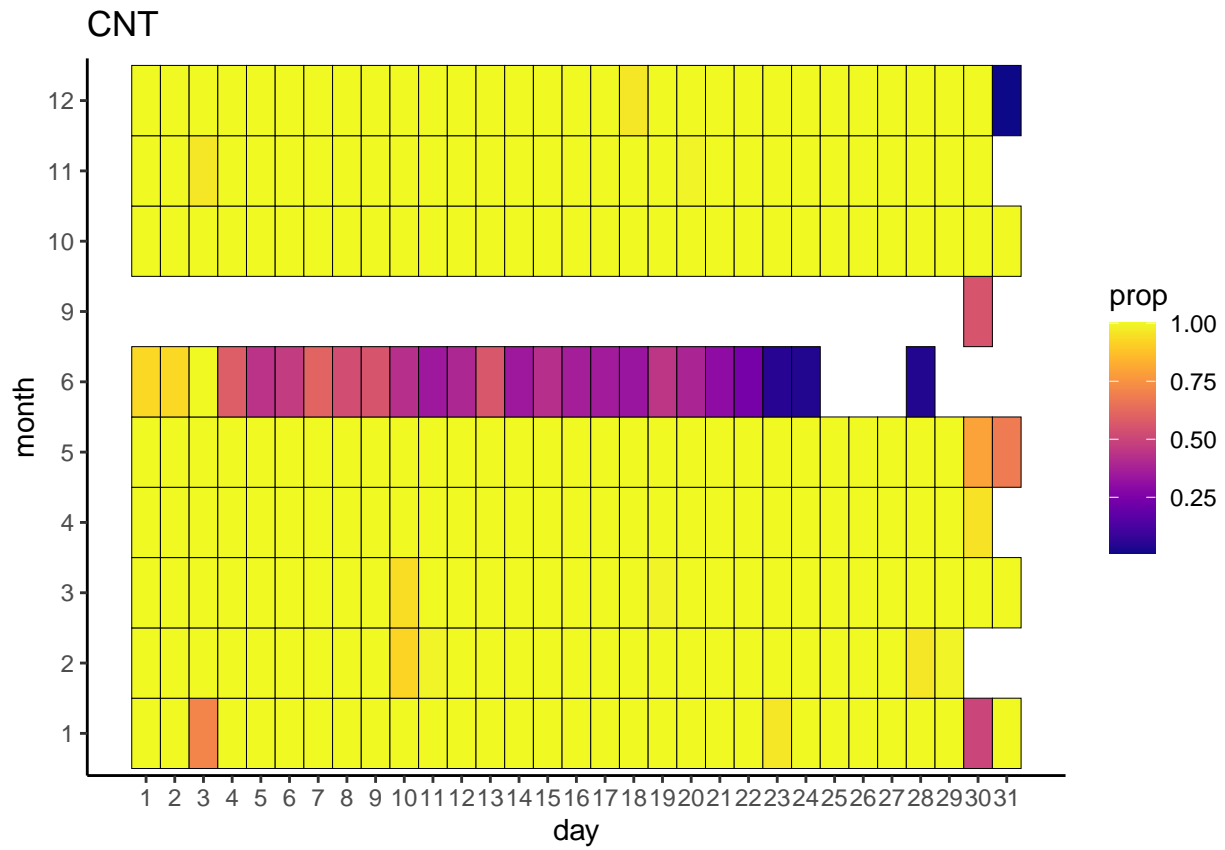
```
plot_ec("CNT")
```


Figure 7: Proportion of EC values collected and within range by day at CNT. White boxes are missing data for the entire day.
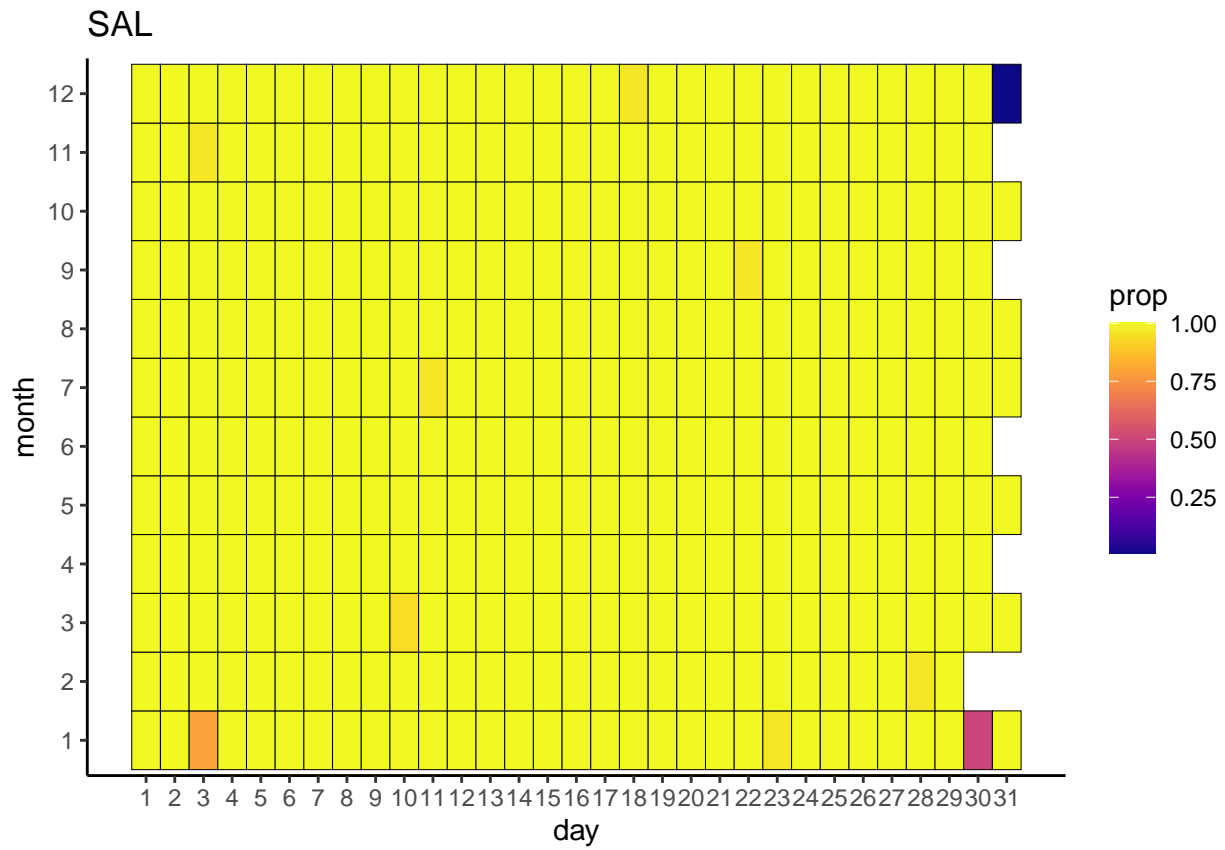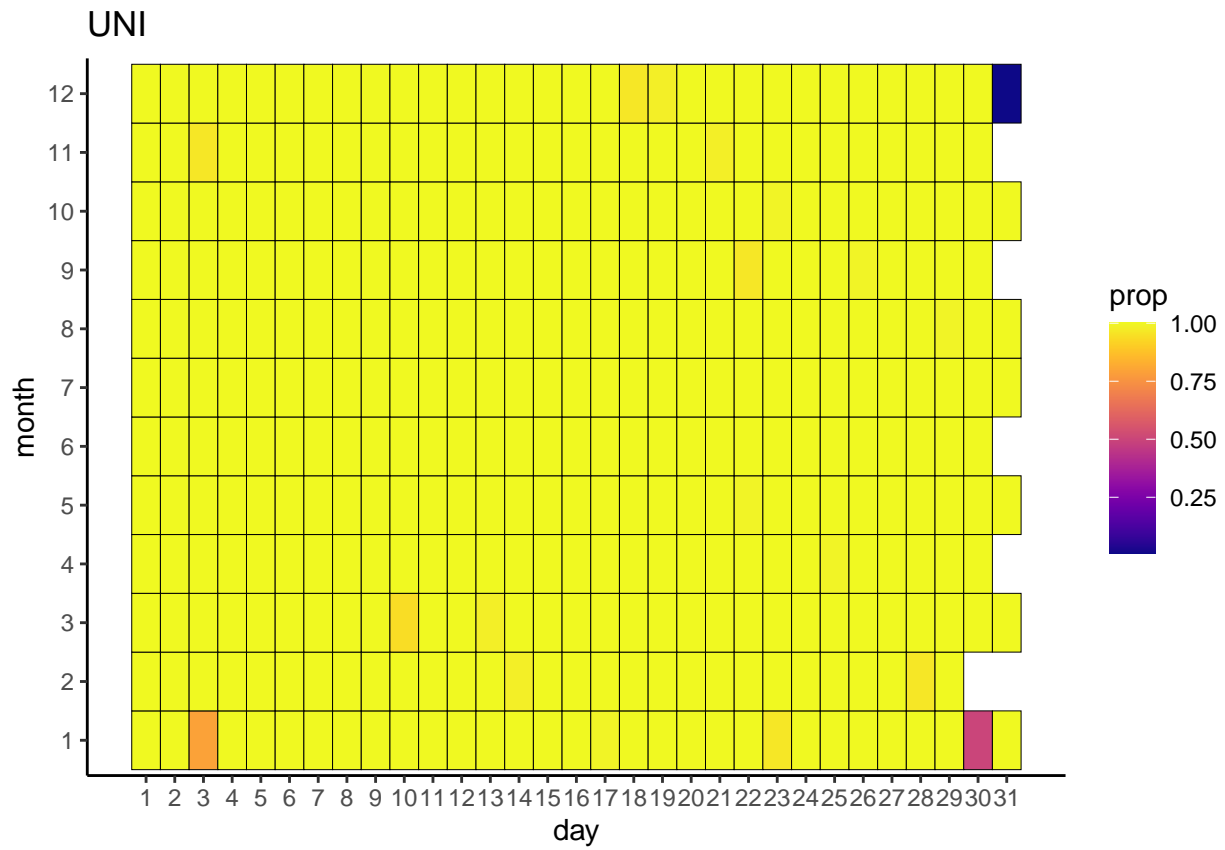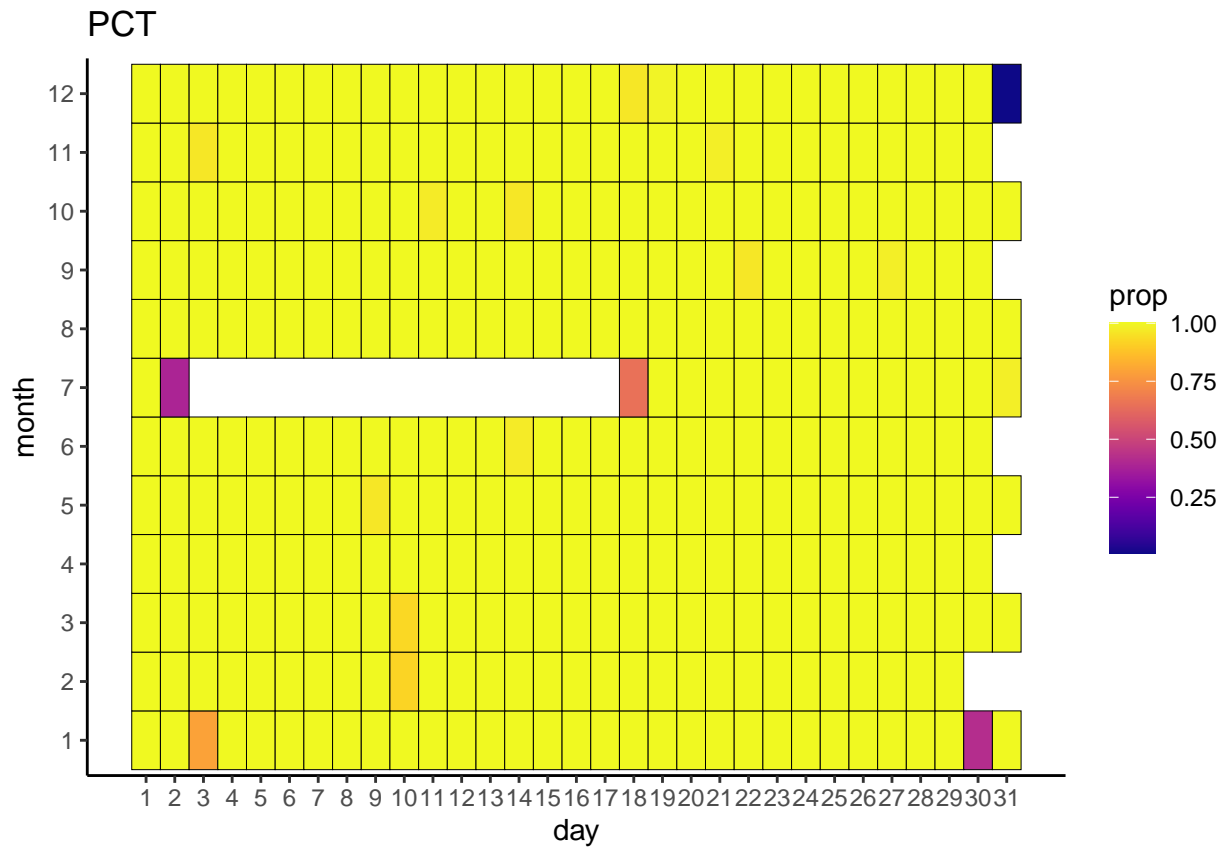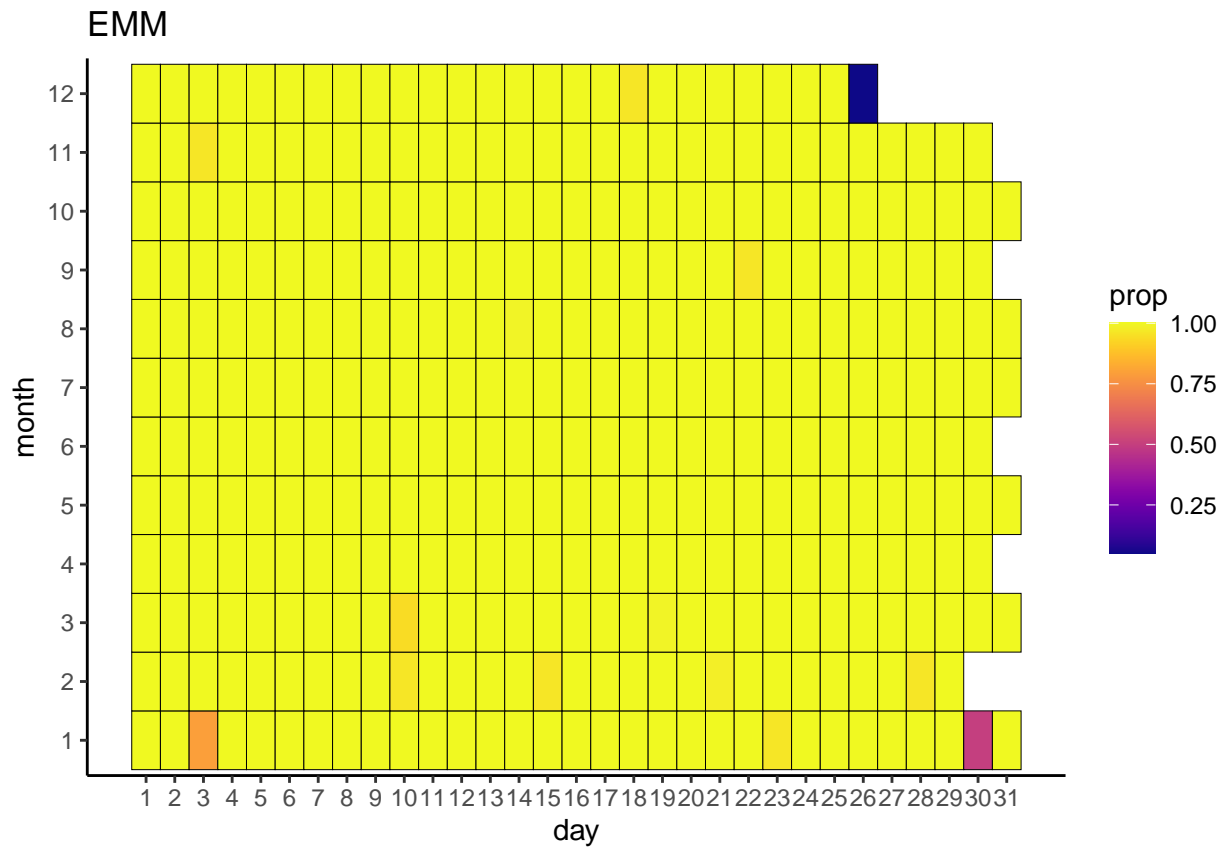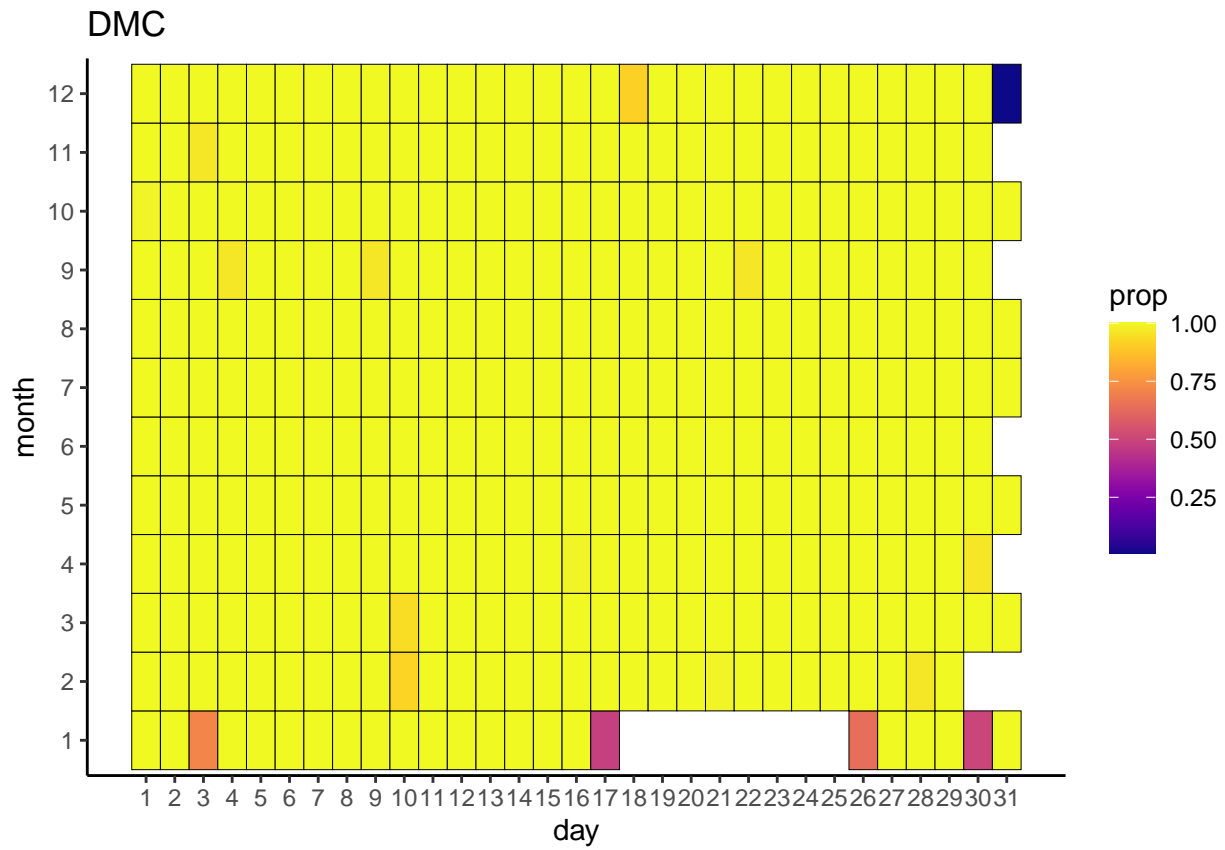
```
plot_ec("SAL")
```


Figure 8: Proportion of EC values collected and within range by day at SAL White boxes are missing data for the entire day.

```
plot_ec("UNI")
```


Figure 9: Proportion of EC values collected and within range by day at UNI White boxes are missing data for the entire day.
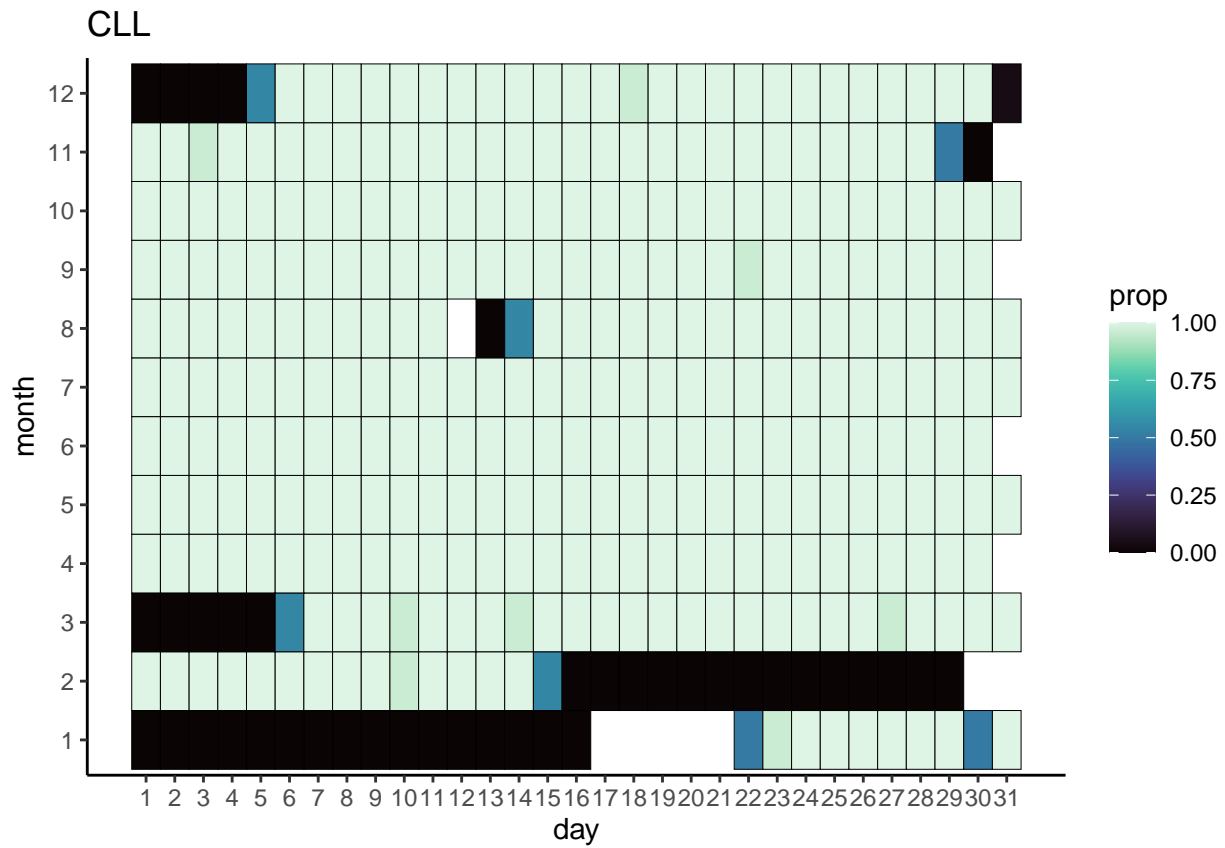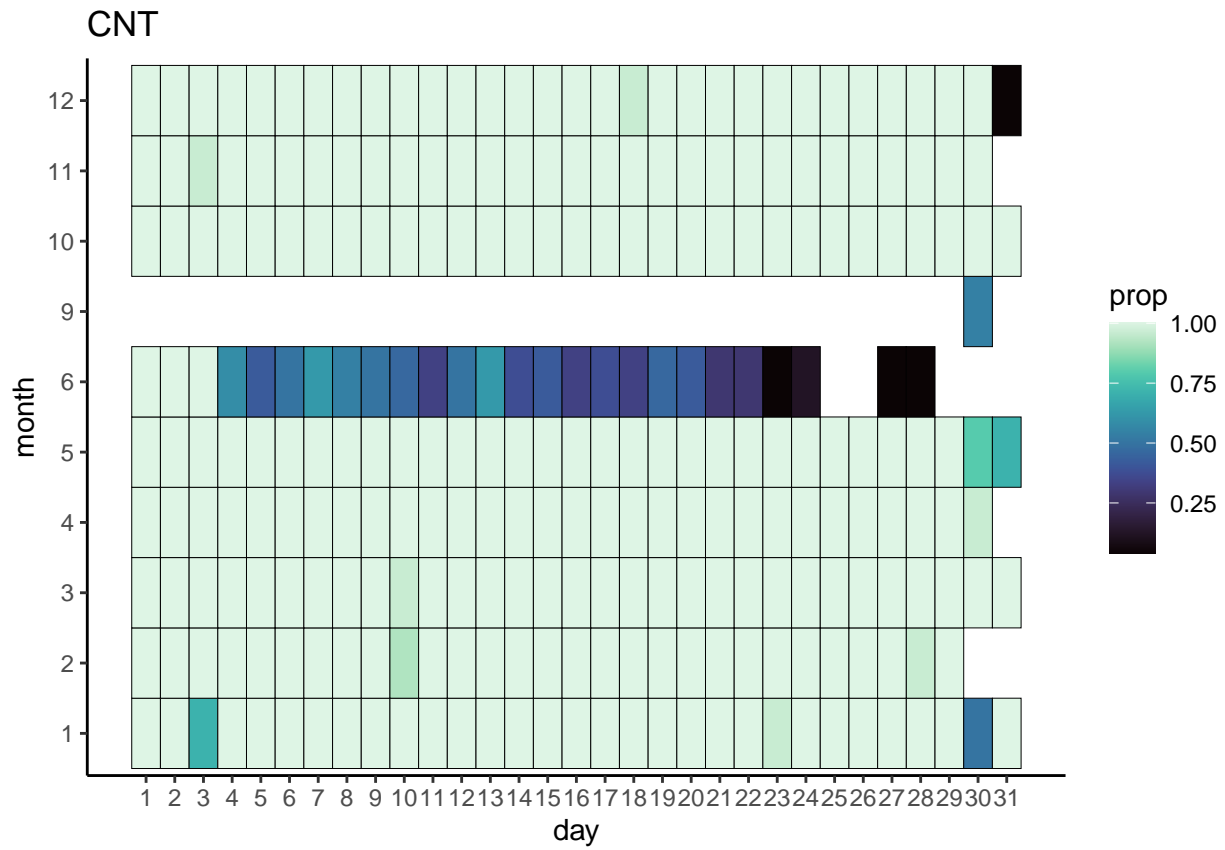
```
plot_ec("PCT")
```


Figure 10: Proportion of EC values collected and within range by day at PCT. White boxes are missing data for the entire day.

```
plot_ec("EMM")
```


Figure 11: Proportion of EC values collected and within range by day at EMM. White boxes are missing data for the entire day.

```
plot_ec("DMC")
```


Figure 12: Proportion of EC values collected and within range by day at DMC. White boxes are missing data for the entire day.

Water Temperature

```
plot_wt("CLL")
```


Figure 13: Proportion of WT values collected and within range by day at CLL. White boxes are missing data for the entire day.

```
plot_wt("CNT")
```


Figure 14: Proportion of WT values collected and within range by day at CNT. White boxes are missing data for the entire day.
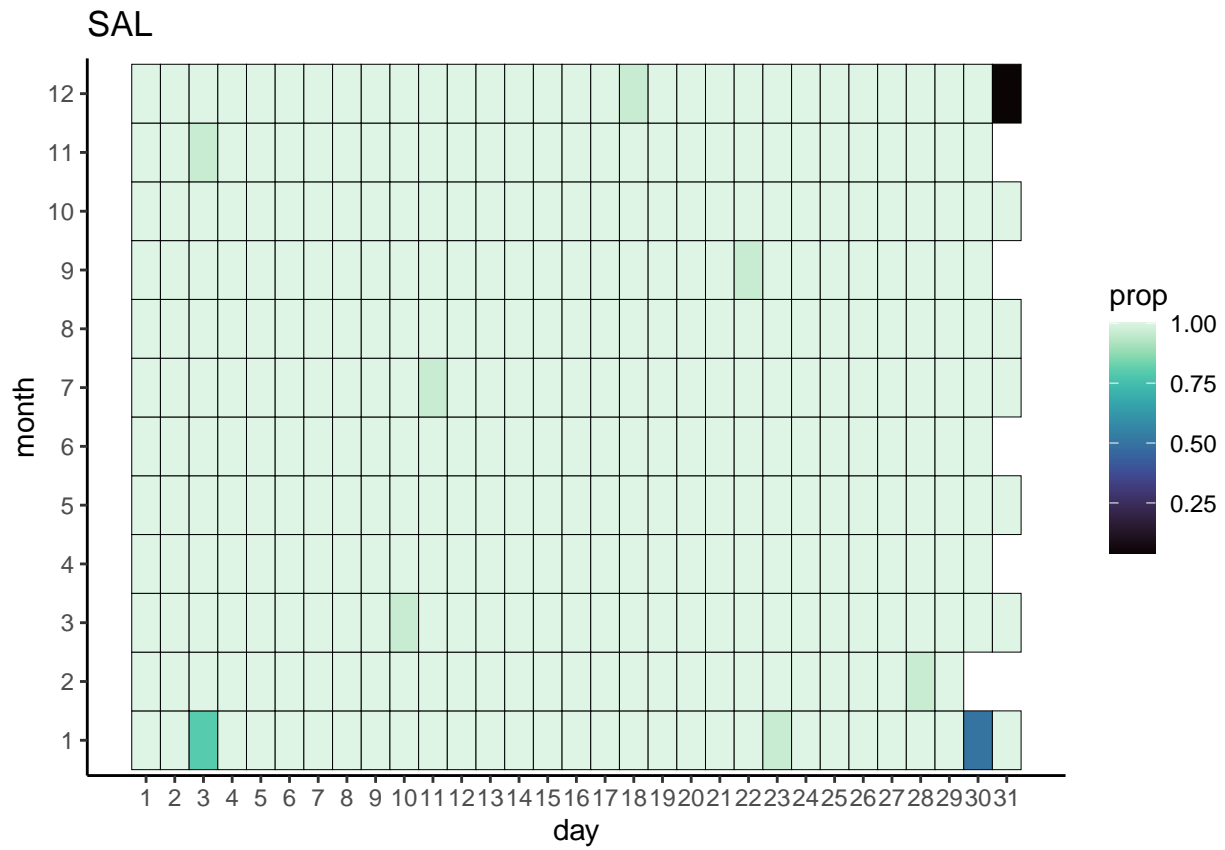
```
plot_wt("SAL")
```


Figure 15: Proportion of WT values collected and within range by day at CNT. White boxes are missing data for the entire day.
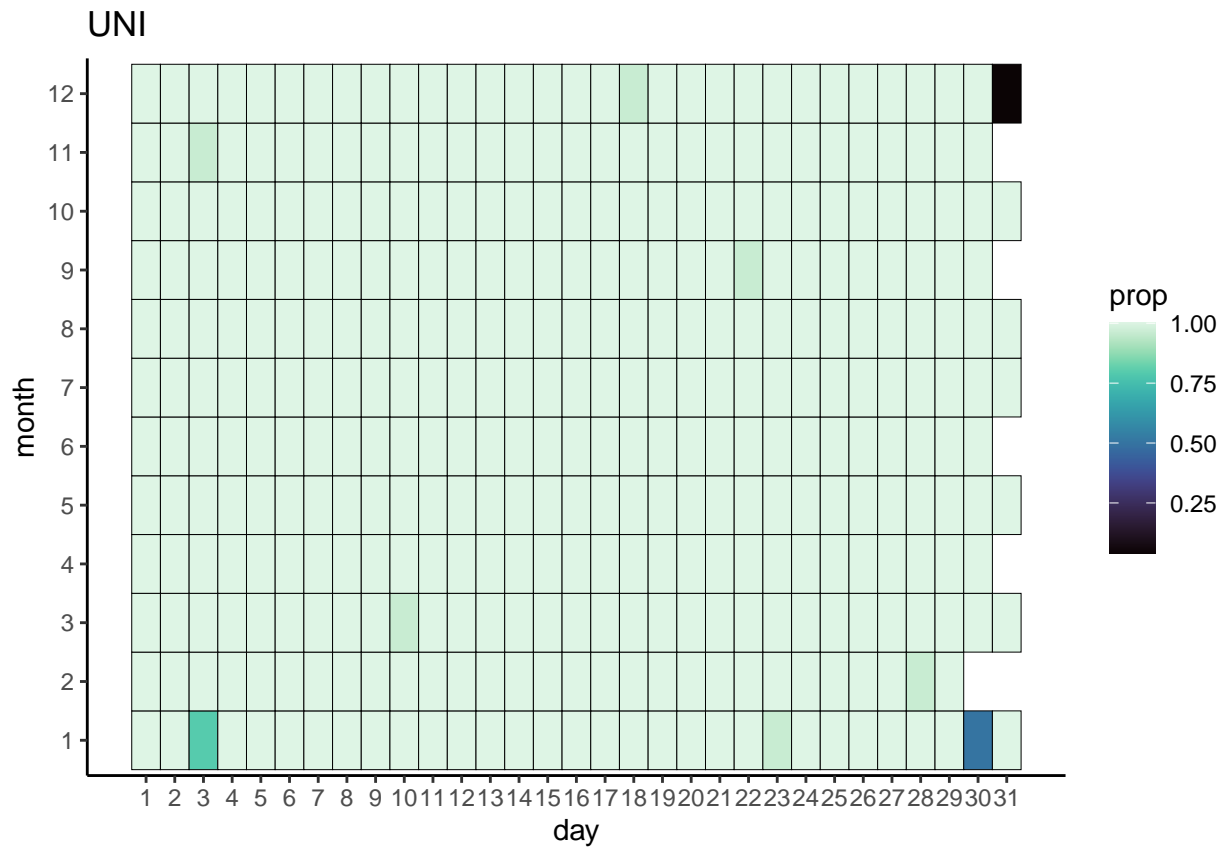
```
plot_wt("UNI")
```


Figure 16: Proportion of WT values collected and within range by day at UNI. White boxes are missing data for the entire day.
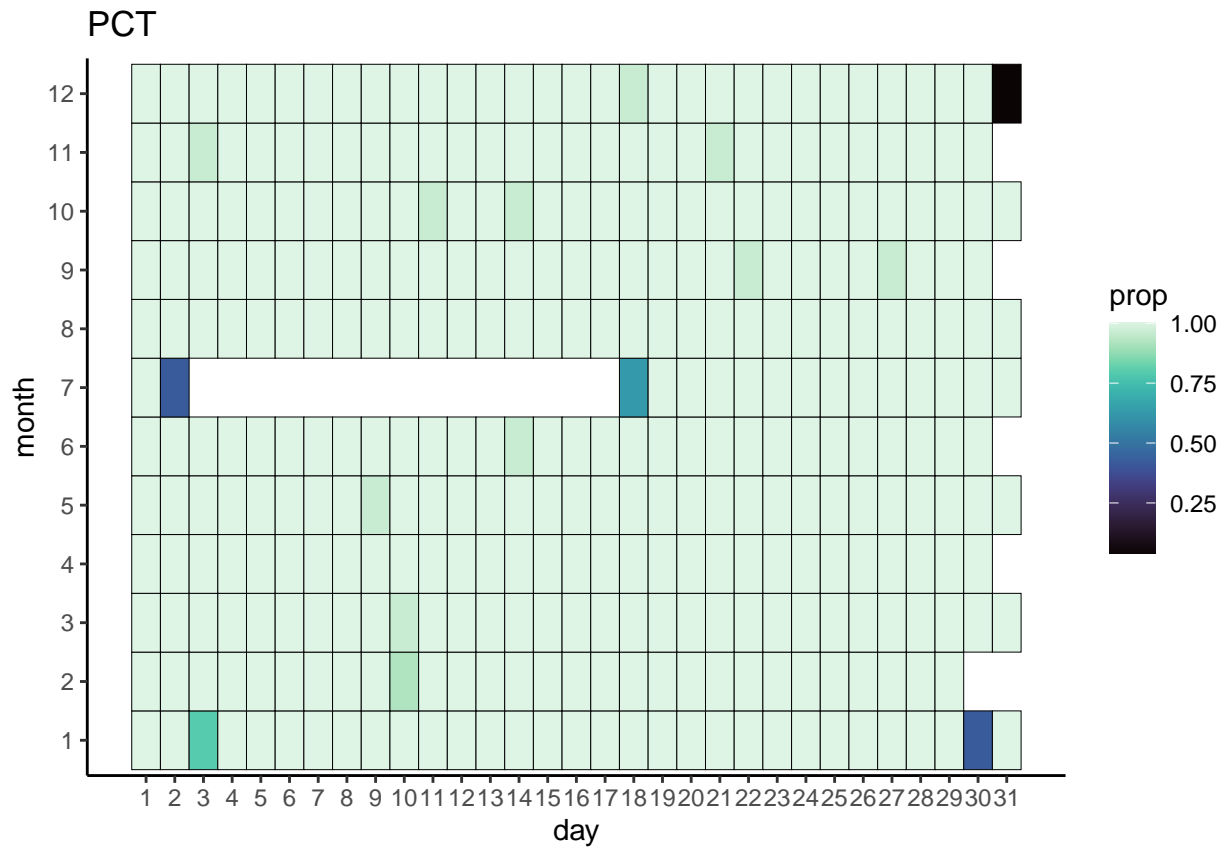
```
plot_wt("PCT")
```


Figure 17: Proportion of WT values collected and within range by day at PCT. White boxes are missing data for the entire day.
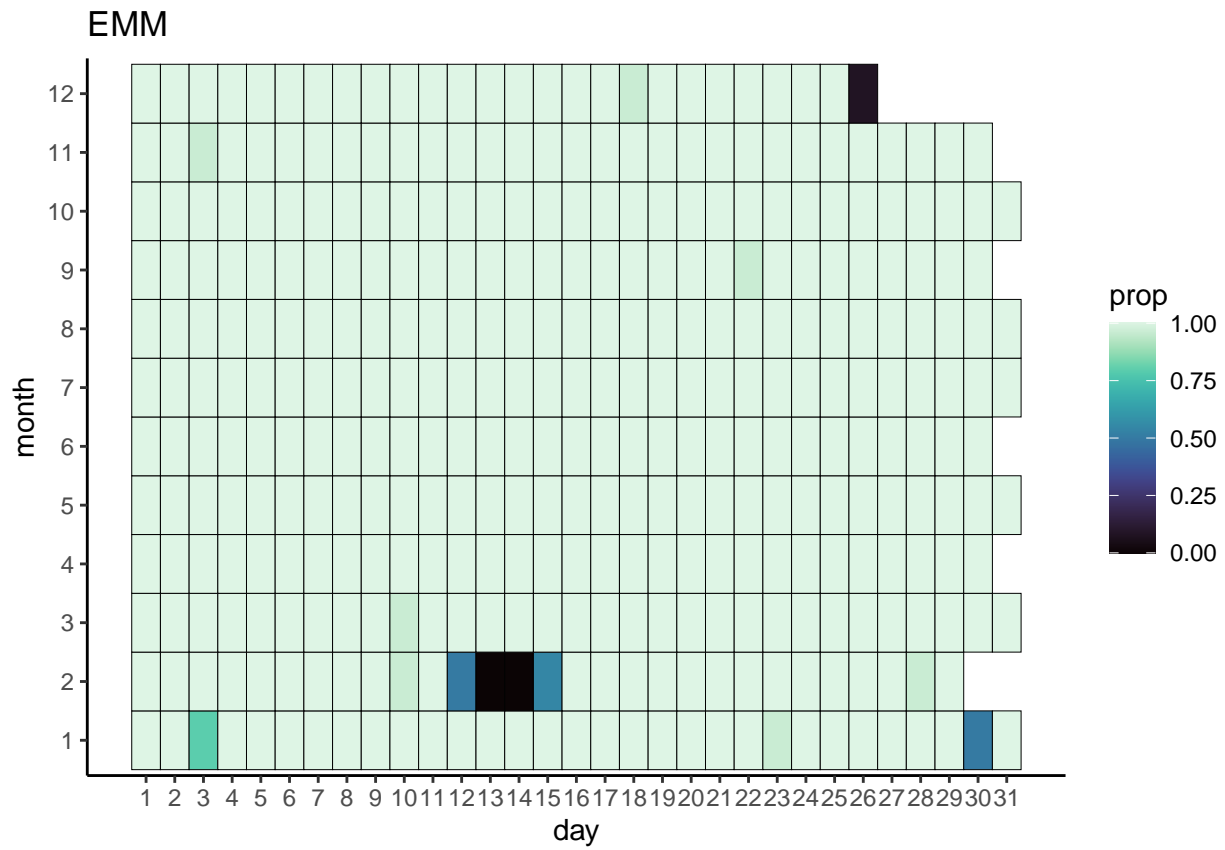
```
plot_wt("EMM")
```


Figure 18: Proportion of WT values collected and within range by day at EMM. White boxes are missing data for the entire day.
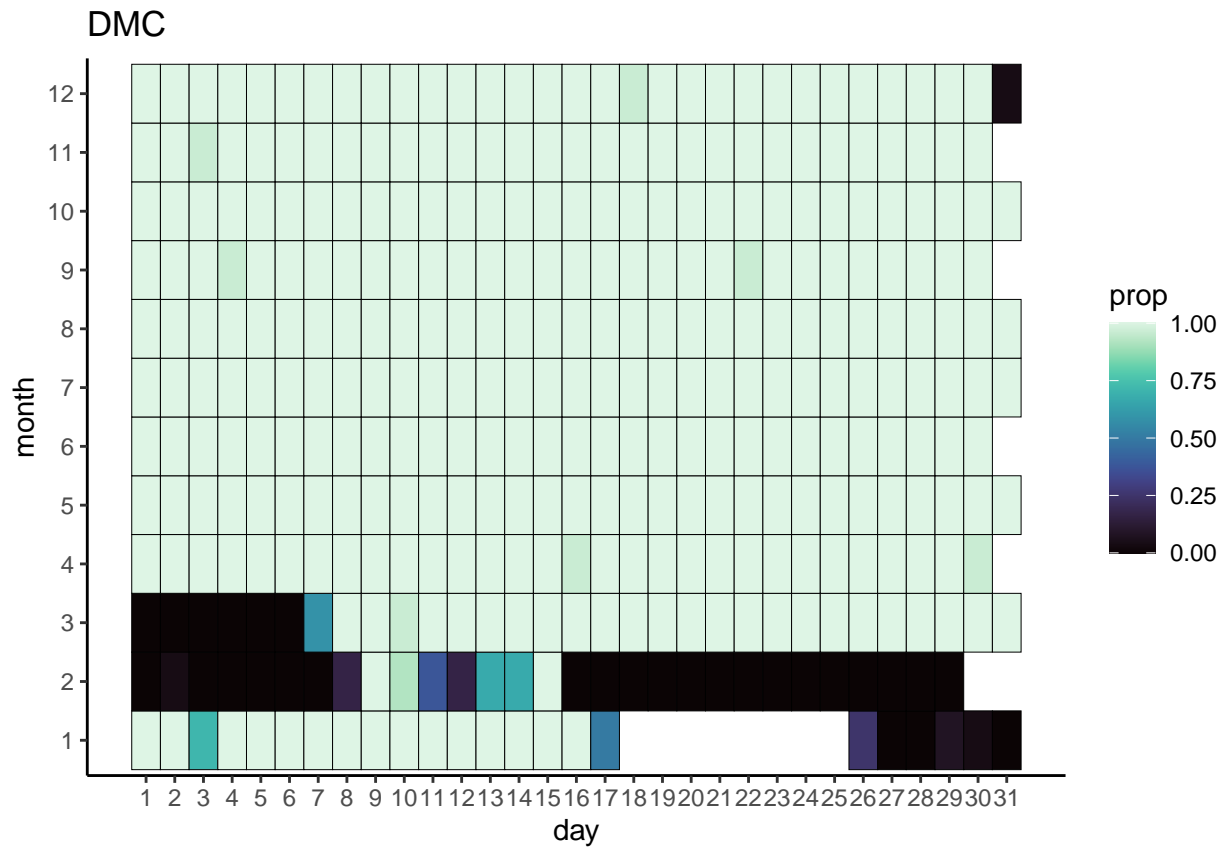
```
plot_wt("DMC")
```


Figure 19: Proportion of WT values collected and within range by day at DMC. White boxes are missing data for the entire day.