Spatial and temporal requirements of breeding Least Terns in the Barataria-Terrebonne National Estuary through the use of Motus automated radio-telemetry data

Zoe Crysler

2019-07-21

Contents

Statistical analysis in Support of Barataria-Terrebonne National Estuary Program Automated VHF Telemetry Network Project

Chapter 1

Methods

1.1 Data Download

Least Tern detection data was downloaded from Motus using the motus R package, and then filtered to remove false positive detections. Detailed information on data download processes, formats, and false detections can be found here: https://motus.org/MotusRBook/.

1.2 Data Cleaning

Strings of detections with only 2 consecutive detections at a station (run length = 2) are questionable and may indicate the tag was just barely within range of the station, or more likely, that it is a false positive produced by background radio noise which is present to varying degrees at all stations.

To begin, several tags were never detected near the breeding site and were removed (2017 Lotek tag ID's 95, 96, 97, 2018 Lotek tag ID 315). Then, detections at stations outside the tagging area with a run length of 2 were deemed to be false positives and removed. Each remaining tag was then visually inspected to examine detections that occurred at questionable latitudes far from the breeding site. Several stations had large numbers of detections with low run lengths, this is characteristic of stations with high levels of background radio noise which can often result in false positives. These stations were *Drasher*, *RTNJ*, *Koffler*, *Waggoners Gap*, *Burntpoint*, *Los Vientos Forest*, and *D'Estimauville*. Further inspection of detections at these stations showed that it would be biologically impossible for birds to make the flight between detections on the breeding ground and subsequent detections at these stations in the amount of time between consecutive detections. Therefore, all detections at these stations were deemed to be false positives due to noisy sites.

Remaining detections far from the tagging location were examined, many had run lengths of only 2 which are questionable at best and were deemed to be false positives. Others with run lengths of 3 or 4 were not biologically possible flights such as between the tagging area and eastern Canada.

All detections deemed to be false positives were added to a custom filter that was applied to the data and were excluded from further analysis.

Detailed data filtering methodology with complete R code is available in Appendix A.

1.3 Data Manipulation

Detailed nest monitoring records were used to determine tag retention and incubation periods. To determine incubation period, each tag was assigned an end date based on the "Fate" variable, once a nest's fate was categorized as either "failed nest", "successful nest", or "eggs due to hatch but no chicks found", it was assigned an end date of that day; the "nestEnd" variable.

Lotek tag ID's 114 and 316 did not have any nest records, as their incubation period could therefore not be determined, they were excluded from further analysis.

Several tags were only detected for a short period of time; for meaningful results, tags that were detected for less than 24 hours were excluded from further analysis (Lotek tag ID's 85, 98, 120, 130, 286, 288, 309, 312).

1.4 Site Visits

To determine the frequency at which Least Terns visited each foraging area, detections at each station were grouped into distinct 'visits'. A visit was categorized as a period of consecutive detections (in minutes) at a single station, or multiple periods of consecutive detections at a single station where there were fewer that 10 minutes between consecutive periods of detection. Detailed visit summaries are available in Appendix C.

For each year, an ANOVA test was initially used to determine if the length of visits differed between stations, results indicated that the mean time spent at stations differed significantly between stations for 2017 ($p < 0.001$) and 2018 ($p = 0.001$). A Tukey Test was then implemented to discern where the difference lay for both years.

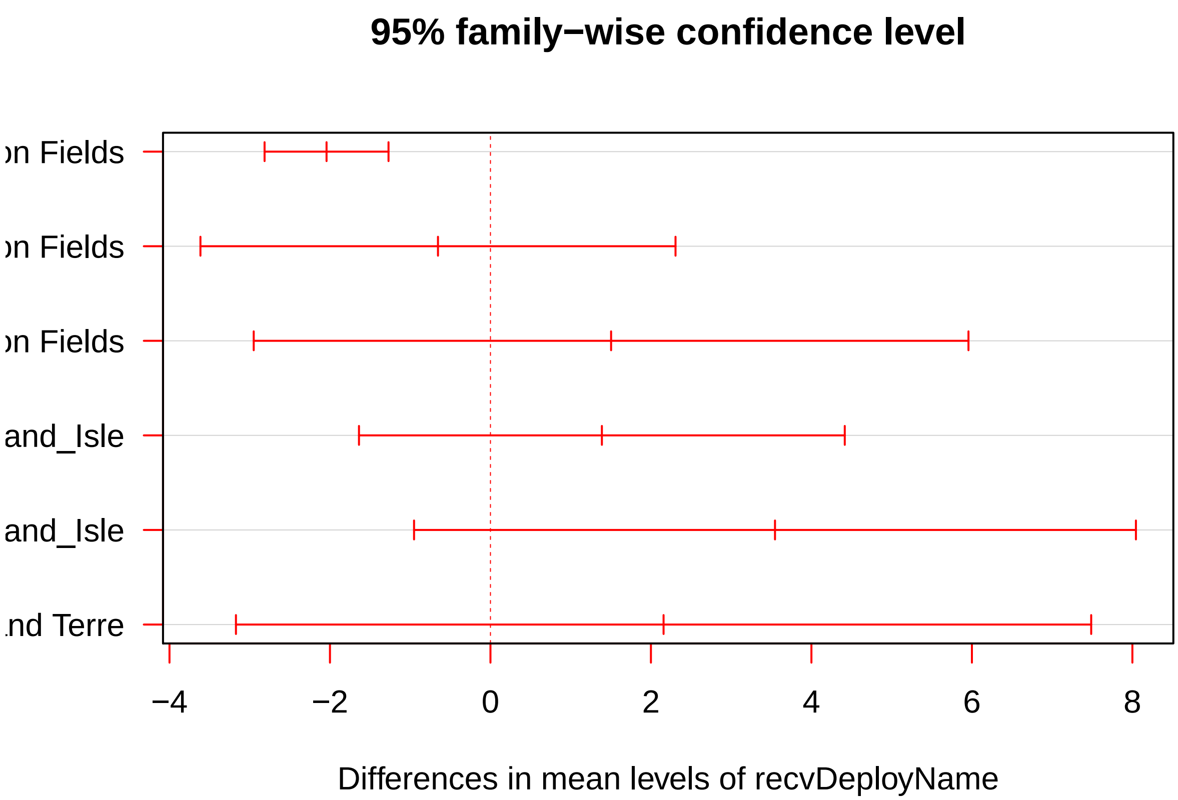


Table 1.1: Incubation detection period for all tags

Year	motusTagID	mfgID	First Detection	Final Detection	Detection Period (days)	Number of Sites
2017	23209	84	2017-06-14 19:13:51	2017-06-16 07:54:11	1.5280160 days	2
2017	23210	85	2017-06-13 18:23:28	2017-06-14 11:49:19	0.7262788 days	2
2017	23212	87	2017-06-13 19:02:47	2017-06-15 08:46:36	1.5720887 days	2
2017	23214	89	2017-06-13 18:45:24	2017-06-18 06:45:36	4.5001392 days	2
2017	23216	91	2017-06-08 15:42:56	2017-06-30 23:53:53	22.3409424 days	3
2017	23217	92	2017-06-08 11:05:46	2017-06-10 06:36:22	1.8129139 days	2
2017	23218	93	2017-06-08 13:11:20	2017-06-19 15:57:39	11.1154951 days	3
2017	23219	94	2017-06-05 17:54:18	2017-06-12 19:34:05	7.0692972 days	2
2017	23223	98	2017-06-08 10:24:45	2017-06-08 12:12:27	0.0747868 days	1
2017	23225	100	2017-06-06 13:42:58	2017-06-11 15:21:11	5.0682081 days	2
2017	23226	101	2017-06-05 16:51:07	2017-06-08 09:21:16	2.6876057 days	2
2017	23227	102	2017-06-13 20:04:12	2017-06-15 06:59:29	1.4550556 days	2
2017	23244	120	2017-05-24 16:28:18	2017-05-24 17:04:05	0.0248418 days	2
2017	23245	121	2017-05-24 17:32:52	2017-05-28 10:13:52	3.6951399 days	2
2017	23246	122	2017-05-24 13:45:29	2017-05-27 09:55:23	2.8402087 days	2
2017	23248	124	2017-05-22 17:48:08	2017-05-25 07:21:54	2.5651166 days	3
2017	23249	126	2017-05-22 16:51:49	2017-05-24 08:48:42	1.6645054 days	2
2017	23251	129	2017-05-22 16:24:28	2017-05-31 09:48:09	8.7247749 days	3
2017	23252	130	2017-05-24 14:26:26	2017-05-24 14:50:43	0.0168660 days	1
2017	23257	135	2017-05-16 15:21:24	2017-05-24 16:44:46	8.0578901 days	3
2017	23258	137	2017-05-19 06:24:57	2017-05-21 16:39:51	2.4270138 days	2
2017	23259	138	2017-05-17 19:10:13	2017-05-22 09:58:09	4.6166171 days	3
2017	23261	141	2017-05-16 08:23:27	2017-05-22 19:57:40	6.4820976 days	3
2017	23262	142	2017-05-16 09:43:35	2017-05-19 08:38:08	2.9545508 days	2
2017	23270	136	2017-05-17 18:38:48	2017-05-22 21:40:44	5.1263457 days	6
2018	28520	283	2018-05-10 14:07:47	2018-05-27 21:27:46	17.3055448 days	3
2018	28521	284	2018-05-10 15:15:55	2018-06-09 10:52:42	29.8172092 days	2
2018	28522	289	2018-05-15 01:45:02	2018-05-18 14:26:29	3.5287893 days	3
2018	28523	290	2018-05-16 11:49:11	2018-05-28 11:44:27	11.9967108 days	3
2018	28524	293	2018-05-16 12:46:21	2018-05-26 23:32:20	10.4486082 days	2
2018	28525	296	2018-05-17 19:21:18	2018-05-27 23:48:20	10.1854387 days	2
2018	28526	306	2018-05-24 18:52:26	2018-05-28 12:10:34	3.7209200 days	2
2018	28527	309	2018-06-08 19:22:32	2018-06-09 00:34:18	0.2164991 days	2
2018	28528	310	2018-06-13 05:40:32	2018-06-14 14:21:22	1.3616923 days	3
2018	28529	313	2018-06-14 23:41:05	2018-06-19 12:46:33	4.5454711 days	3
2018	28592	298	2018-05-22 18:22:33	2018-05-26 08:53:39	3.6049341 days	1
2018	28593	303	2018-05-24 01:55:54	2018-05-28 12:06:23	4.4239472 days	2
2018	28595	282	2018-05-10 12:00:41	2018-05-21 12:32:49	11.0223161 days	3
2018	28597	286	2018-05-14 11:53:26	2018-05-14 23:19:34	0.4764758 days	2
2018	28598	287	2018-05-14 13:37:52	2018-05-18 07:44:11	3.7543804 days	2
2018	28599	288	2018-05-14 13:41:57	2018-05-14 21:58:19	0.3446921 days	2
2018	28600	291	2018-05-14 19:10:30	2018-05-19 19:28:19	5.0123684 days	3
2018	28602	294	2018-05-16 16:35:22	2018-05-27 11:50:49	10.8024021 days	3
2018	28603	295	2018-05-18 06:01:36	2018-05-26 15:55:30	8.4124330 days	2
2018	28604	297	2018-05-17 19:03:49	2018-05-28 23:26:21	11.1823128 days	2
2018	28605	299	2018-05-22 13:04:44	2018-05-31 17:47:37	9.1964506 days	3
2018	28606	300	2018-05-22 17:37:26	2018-05-27 08:19:37	4.6126208 days	2
2018	28607	301	2018-05-22 23:02:03	2018-05-27 23:59:39	5.0399898 days	2
2018	28608	302	2018-05-22 19:32:17	2018-05-28 14:22:10	5.7846344 days	3
2018	28609	304	2018-05-23 23:56:37	2018-05-28 11:26:27	4.4790540 days	2
2018	28611	307	2018-06-08 23:06:12	2018-06-25 18:59:41	16.8288073 days	3
2018	28613	312	2018-06-14 16:36:05	2018-06-14 21:49:50	0.2178772 days	3

Chapter 2

Results

2.1 Study Area

Stations within the study region are presented in Figure 2.1, stations with red outlines had no detections, while coloured stations had at least one tag detected during the study period from 2017 to 2018. Triangles represent receivers that were active during 2017 *and* 2018, circles represent stations that were only active in 2017 (*Quintana* and *East Grand Terre*). The blue lines indicate the direction of antennas during the study period and span 15 miles from the receiver. All receivers maintained the same bearings throughout the study period with the exception of *Bolivar_Flats*, here the solid blue lines are the 2017 bearings, while the dashed blue lines are the 2018 bearings.

All Least Terns were tagged within close proximity of *Exxon Fields*, and for the purposes of this report stations considered within proximity to the breeding area are *Exxon Fields*, *East Grand Terre*, *Grand_Isle*, and *Wisner.* A close up of this area is presented in the inset of figure 2.1

2.2 Detection Overview

In 2017, only one individual (Motus tag ID 23270) was detected at stations outside the immediate tagging region, all other individuals and the bulk of the detections occurred at *Exxon Fields*, *East Grand Terre*, *Grand_Isle*, and *Wisner.* Figure 2.2 shows a map of all 2017 tagged bird detections, lines are coloured by tag ID. Figure 2.3 displays the same data by plotting latitude vs. time for each tag, with points coloured by receiver site.

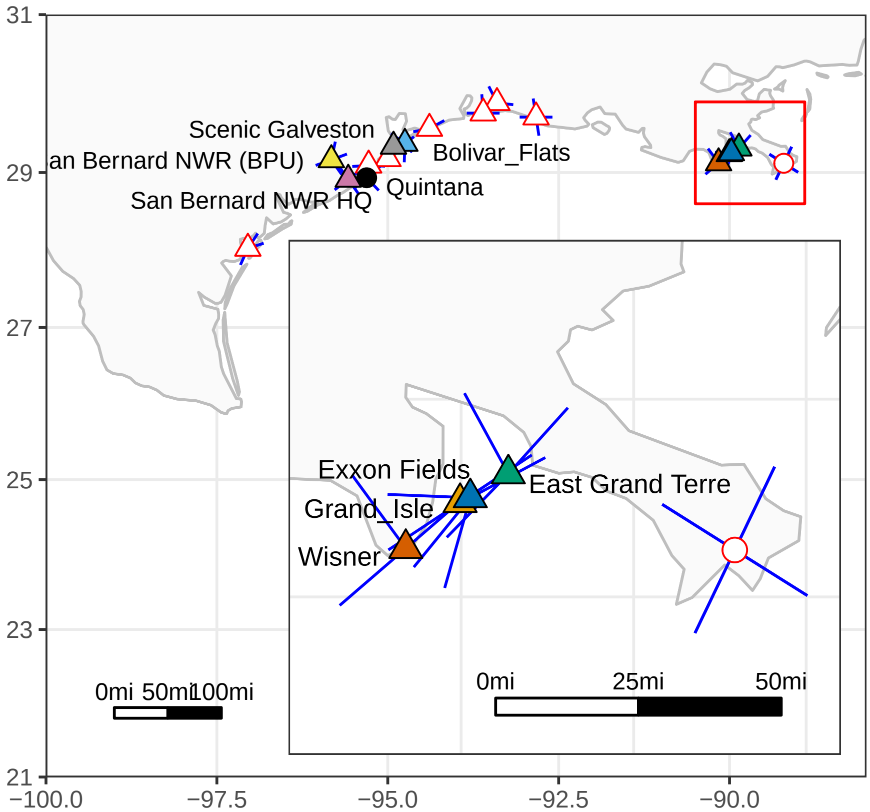
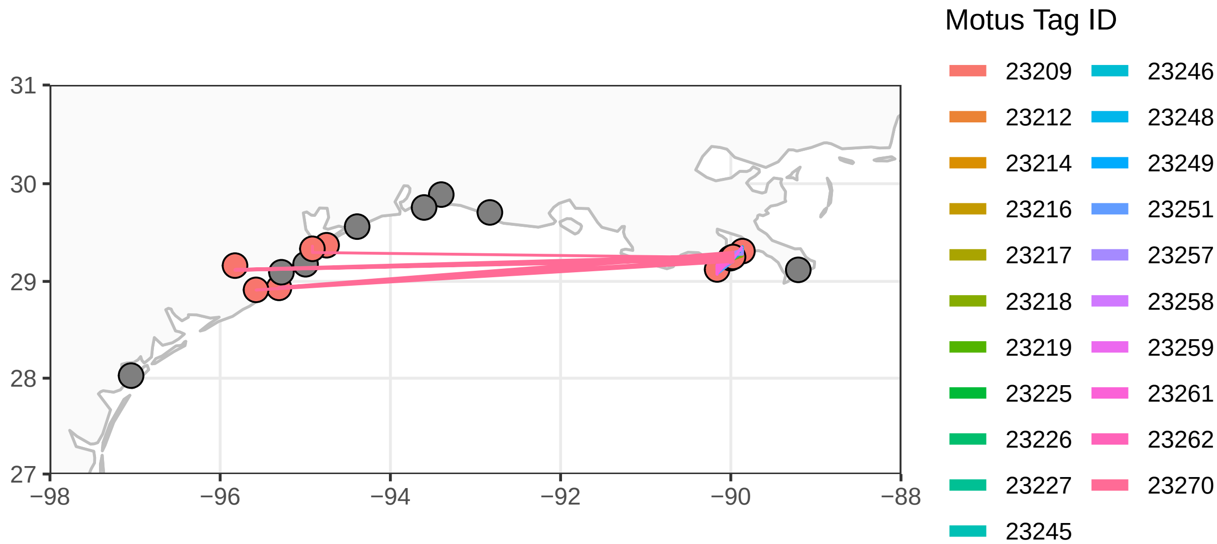


Figure 2.1: Map of study region during 2017 and 2018. Triangles represent receivers that were active during 2017 and 2018, circles represent stations that were only active in 2017. Stations outlined in a red circle had no detections while coloured stations had at least one detection during the study period. Blue lines indicate antenna bearings spanning 15 miles from the station, bearings remained constant throughout the study period with the exception of *Bolivar_Flats* where the solid lines indicate 2017 bearings and dashed lines indicate 2018 bearings.



In 2018, again only one individual was detected outside the immediate tagging region (Motus tag ID 28520), all other individuals and the bulk of the detections occurred at *Exxon Fields*, *East Grand Terre*, and *Grand_Isle*. Note that in 2018 no birds were detected at *Wisner* despite close proximity to nesting areas. Based on receiver performance outputs, this is likely due to receiver functionality rather than a true lack of detections. Figure 2.4 shows a map of detection pattern of all 2018 tagged birds, while figure 2.5 displays the same data as individual tag plots of latitude vs. time.

2.3 Foraging vs Incubating

Figure 2.6 is a selection of 4 representative detection patterns of individuals; examining the signal strength it is not always clear when birds are attending the nest site and when they are foraging. For example, tag 23219 shows clear daily signal strength patterns, with low signal strength occurring overnight when the bird is likely incubating. Tag 23218 has a similar but less distinct pattern; some nights the bird is detected while other nights it is not. In other cases the tag is picked up far less consistently as in tag 28593, and in even more extreme cases as in tag 28522. For these final two tags, the bird is rarely if ever detected while incubating, and most detections likely occur as the bird is leaving or arriving at the nest site.

Furthermore, even tags with clear daily detection patterns as in tag 23219, shorter nest site attendance periods during the day are not as obviously picked up. However, based on these detections, the signal strength during incubation is generally much lower than during flight. This is to be expected as a tags detection range increases substantially when a bird is in the air as compared to on the ground (low signal strength = weaker detection). Therefore, all detections on the Exxon Fields receiver with a signal strength < 80 will be excluded from foraging analysis from this point forward to exclude likely detections on the nest during incubation. While this will invariably also remove detections on Exxon Fields that may occur when the bird is in flight, since the station is in very close proximity to the nesting site, in most cases this will only remove the tail

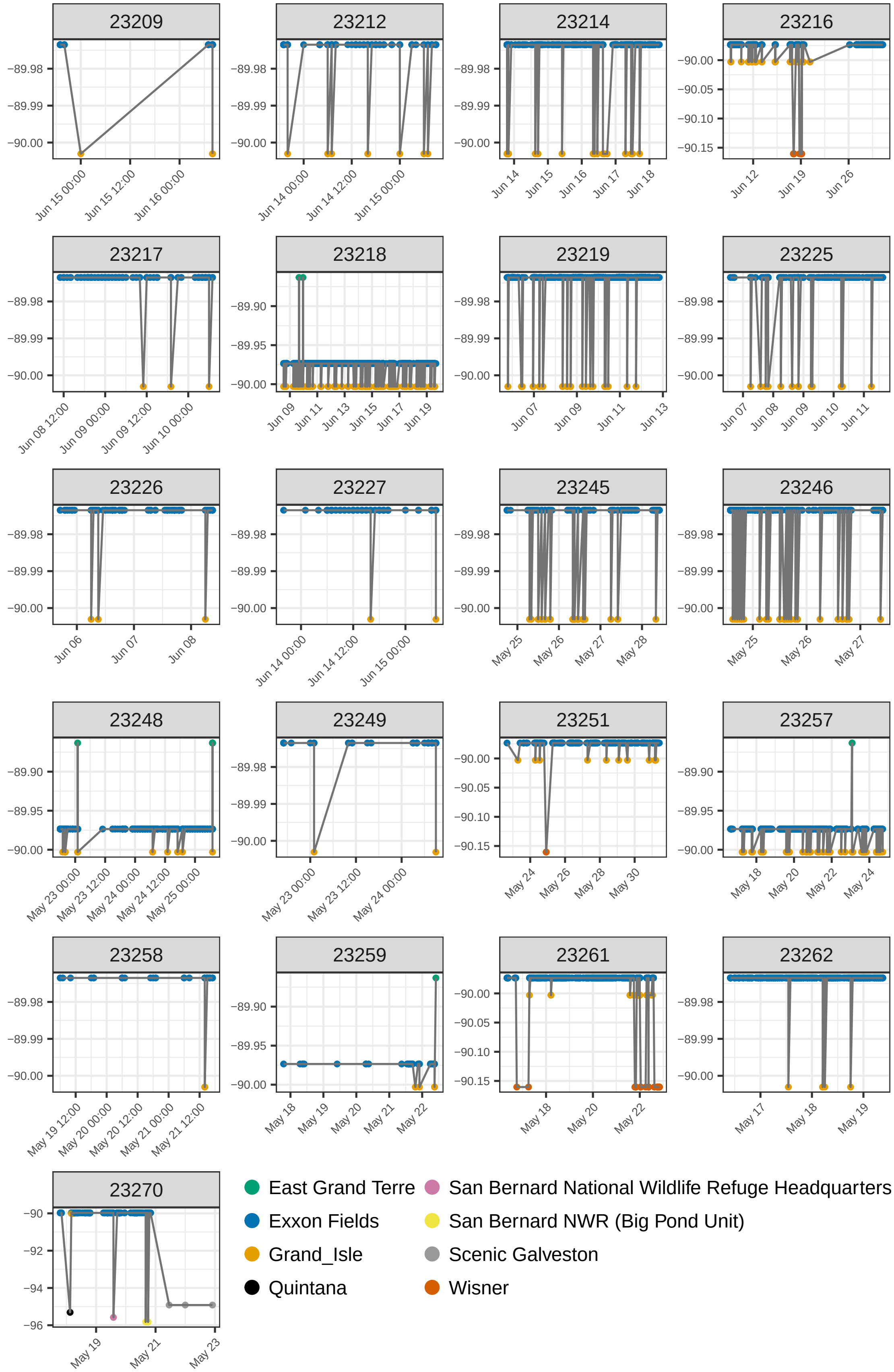


Figure 2.2: Individual 2017 plots by Motus tag ID showing latitudinal movements of tags over time, detection points are coloured by station location. Note that the x and y scales vary between tag plots.

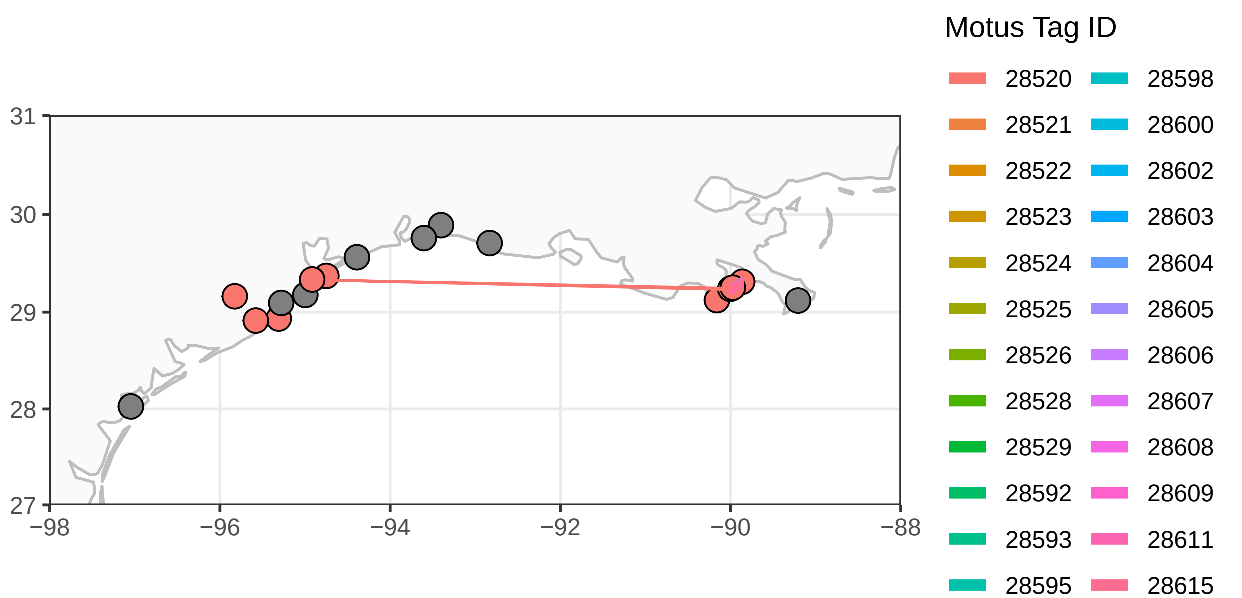


Figure 2.3: Map of 2018 tag detections, only tag 28520 was detected outside the immediate tagging region.

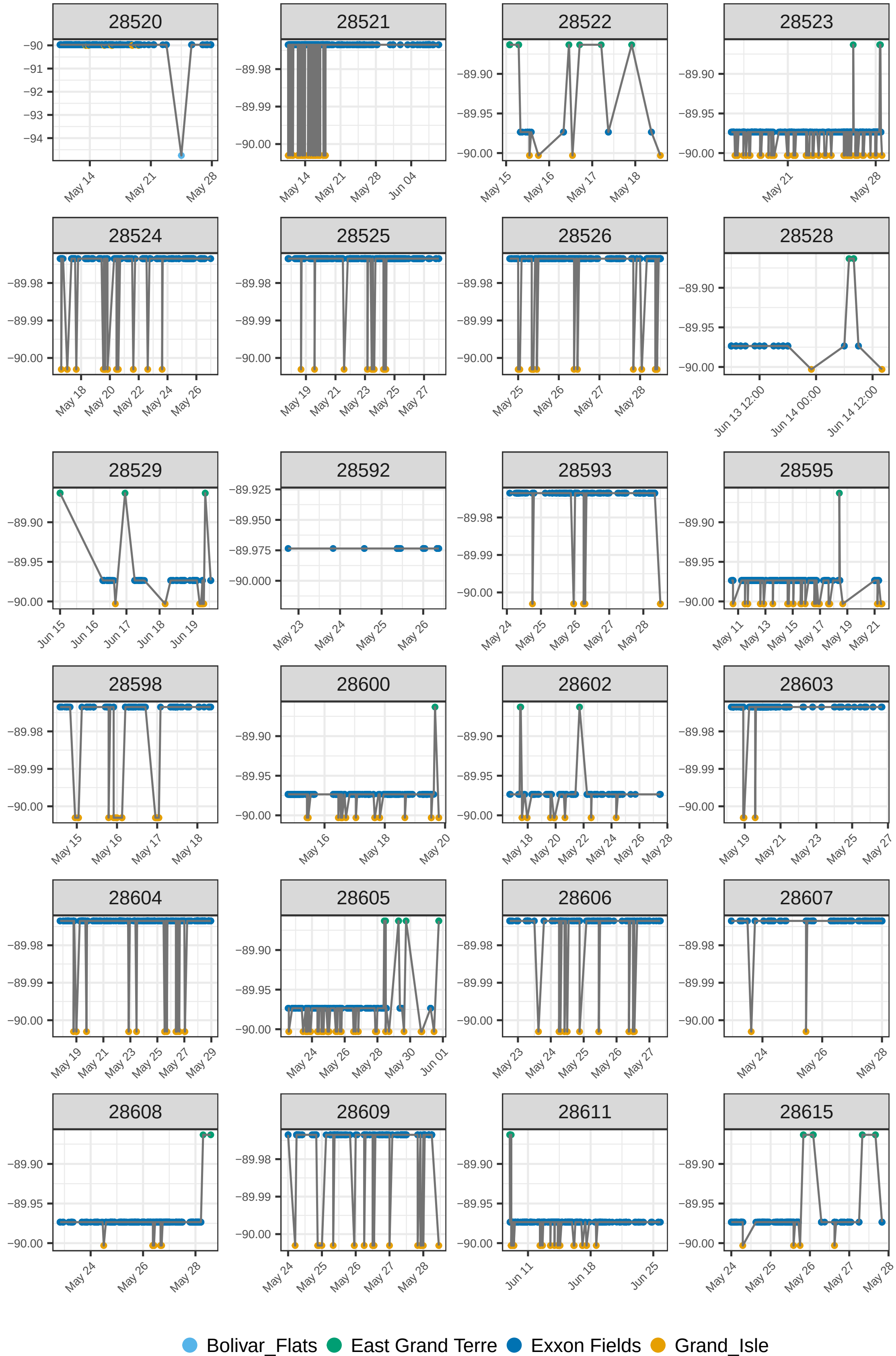


Figure 2.4: Individual 2018 plots by tag ID showing latitudinal movements of tags over time, detection points are coloured by station location. Note that the x and y scales vary between tag plots.

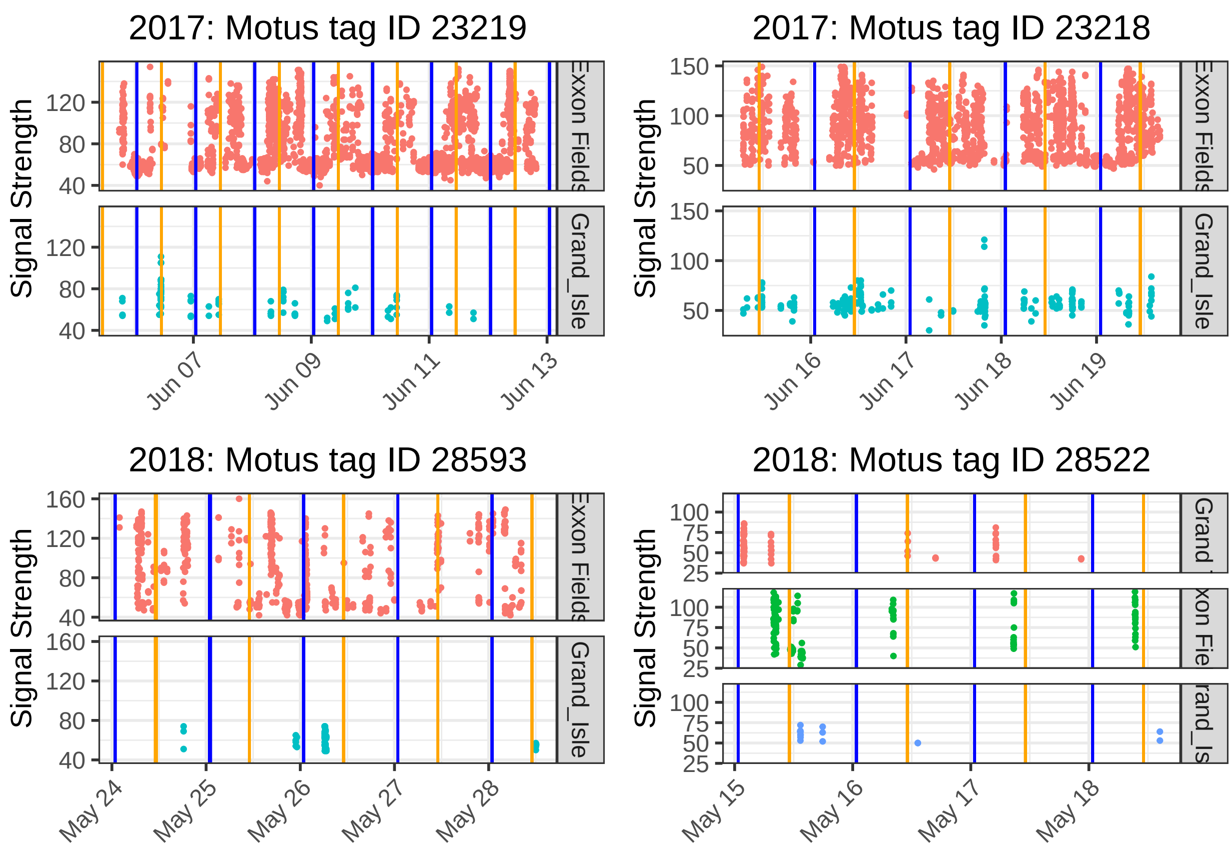


Figure 2.5: Signal strength vs. time of four representative tags at select periods of incubation. Points represent individual detections coloured by Receiver, vertical blue lines indicate sunset, and vertical orange lines indicate sunrise. A higher signal strength correlates to a stronger detection. Note x and y scales vary between plots.

Table 2.1: Distance in miles between each receiver and Exxon Fields

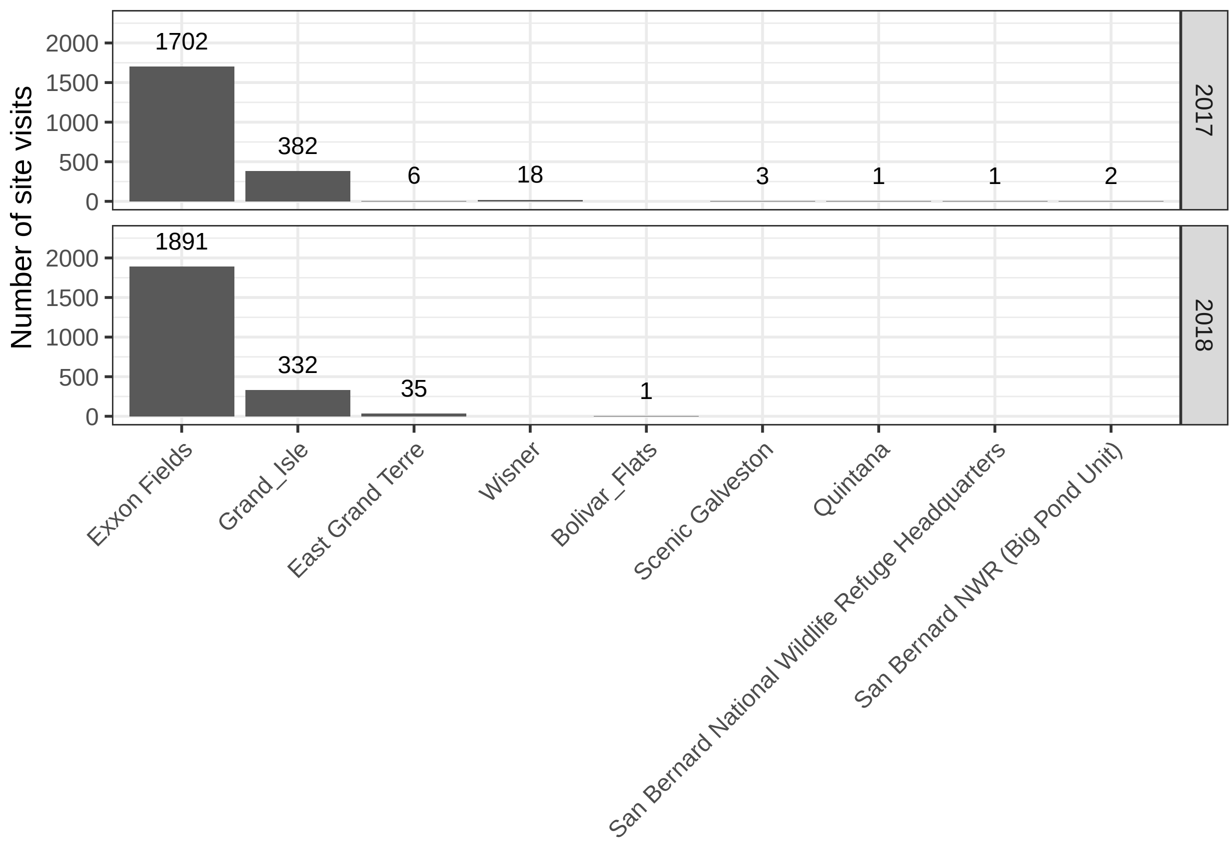
Site Name	Distance from Exxon Fields (mi)
Exxon Fields	0.000000
Grand_Isle	1.967003
East Grand Terre	7.848415
Wisner	14.336358
Rock W	175.022123
Sabine NWR	211.054814
High_Island	267.210775
Bolivar_Flats	288.394805
Scenic Galveston	298.436099
Brazoria National Wildlife Refuge	320.902090
Quintana	323.455637
San Bernard National Wildlife Refuge Headquarters	339.849524
San Bernard NWR (Big Pond Unit)	353.617184
Rockport Bay Education Center	437.851102

ends of a detection streak as the bird is departing or approaching the nesting site.

2.4 Heavily Frequented Foraging Areas

To add meaning to receiver detections, the distance from *Exxon Fields* was calculated for each station where tags were detected and are presented in table 2.1. A full database with distances calculated between all stations can be found in Appendix B.

For each station the total number of visits was calculated (figure 2.8), as expected *Exxon Fields* had by far the highest number of visits as it was at the nesting area. In 2017 and 2018 *Grand_Isle* had the next highest frequency of detections, this is not surprising as it is the closest station to the nesting area (see table 2.1). *East Grand Terre* and *Wisner* followed with lower numbers of site visits and are also near the nesting area. Note the lack of detections at *Wisner* in 2018, this is likely due to receiver malfunction rather than a true lack of detections. All other stations had negligible (<3) visits and are substantially farther away from the nesting grounds.



The length of visits was consistent with the number of visits, that is, birds generally had longer detection periods at *Exxon Fields*, followed by *East Grand Terre*, *Grand_Isle*, and *Wisner*. However, for both 2017 and 2018 only *Grand_Isle* and *Exxon Fields* differed significantly ($p < 0.001$) in their visit length. Figure 2.8 shows the mean length of time each bird was detected at each station during each visit (See Appendix C for complete visit information).

Of the birds that were detected at *Exxon Fields* in 2017 and 2018, all were detected on at least 6 different visits. At *Grand_Isle*, the majority (39/44, 88.6%) of birds were detected on more than 2 separate occasions (>2 discrete visits). In comparison, at the remaining stations less than half (10/23, 43.5%) were detected on more than two occasions. Only tag 23261 appeared semi-regularly at *Wisner* but is notable in 12 detections.

Thus birds appear to spend the majority of their time in close proximity to *Exxon Fields*, however this could be a by-product of increased detection due to proximity to nesting sites and therefore a necessity to frequently pass the receiver regardless if foraging is occurring nearby or not. Most birds frequented *Grand_Isle* regularly which could be due to nearby foraging, or a proximity to the nesting site.

2.4.1 Simultaneous Detections

If stations are within close proximity of one another, it's possible to get simultaneous detections on both stations at once. During incubation there were 161 pairs of simultaneous detections, 2 of these occurred between *East Grand Terre* and *Exxon Fields*, 3 between *East Grand Terre* and *Grand_Isle*, and the other 156 between *Exxon Fields* and *Grand_Isle*. *Exxon Fields* and *Grand_Isle* are the closest two stations at 1.97 miles apart, easily within range of each other particularly when birds are in flight (see Appendix B). *East Grand Terre* is 7.85 miles and 9.80 miles away from *Exxon Fields* and *Grand_Isle* respectively, the next closest pairs of stations.

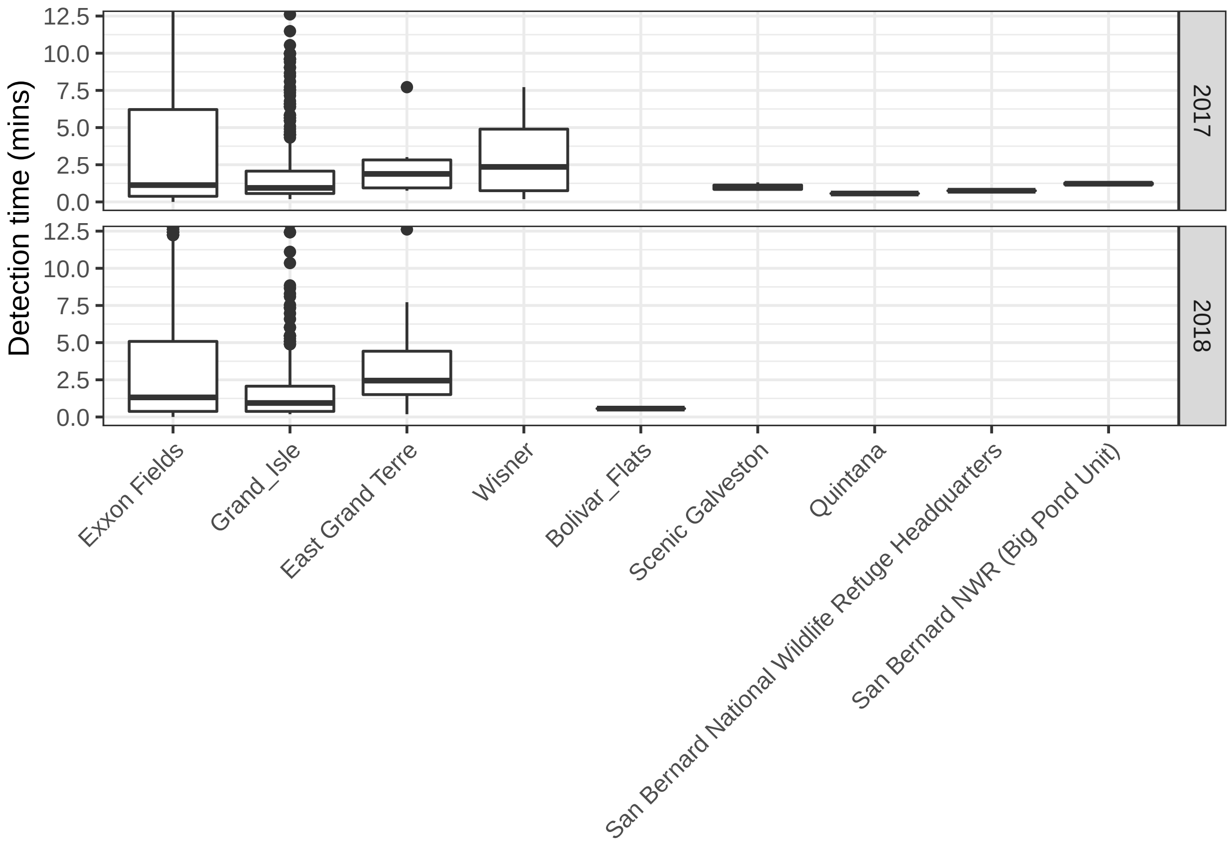


Figure 2.6: Mean length of foraging visits in minutes at each station for 2017 and 2018, ordered by distance from Exxon Fields. Note that the plot has been cut off due to the high number of outliers, *Exxon Fields* had a maximum visit length of 61 minutes, *Grand_Isle* 107 minutes, *East Grand Terre* 19 minutes, *Wisner* 27 minutes.

2.4.2 Detection Directions

We can narrow down the areas visited beyond which site they were detected at by examining which antenna they were detected on. Figures 2.7 and 2.8 show the total number of detections for each antenna bearing at each site for 2017 and 2018 respectively. Orange lines indicate the antenna bearings at each site, the blue bars are coloured by the number of detections on each antenna throughout the incubation period - note that scales differ by site.

Detection patterns are similar at *Exxon Fields*, *Grand_Isle*, and *East Grand Terre* in 2017 and 2018. Detections at *Exxon Fields* were predominantly to the SW, directly in line with the tagging area, this also indicates that birds likely arrive and depart to/from the nesting location from the SW. *Grand_Isle* detections were overwhelmingly more prevalent to the NE, pointing towards the nearby colony, while this may indicate that birds don't spend much time foraging to the SW, it could also be due to the sheer proximity of the nesting site. As discussed earlier, nearly all cases of simultaneous detections between two stations occur between *Exxon Fields* and *Grand Isle*, suggesting that many of the arrival and departure flights from the colony could be captured by both stations, resulting in the high number of NE detections at Grand Isle being an artifact of colony attendance. *East Grand Terre* had the most detections by far to the SW, again towards the colony.

In 2017, birds were picked up in very small numbers outside of the immediate breeding area. It is difficult to make generalized assumptions from such a small number of detections, however it does seem as though a few birds may be using the areas off the coast around *Quintana*. There are even fewer detections away from the colony during incubation in 2018, but those that do occur are in the same region.

2.5 Mean Distances Travelled for Foraging

Based on detection data, in 2017 the mean distance travelled for all detection periods was 1.57 miles, with a maximum distance travelled to *San Bernard NWR (Big Pond Unit)* at 353.62 miles. In 2018 the mean distance travelled was 0.54 miles, with a maximum distance travelled of 288.39 miles to *Bolivar Flats*. However, the vast majority of detections (4,366/4,374 or 99.8%) during foraging trips occurred within 14.34 miles from the colony detected at the four closest receiver stations (*Exxon Fields*, *Grand_Isle*, *East Grand Terre*, and *Wisner*).

2.6 Time Partitioning Between Foraging and Nest Site Attendance

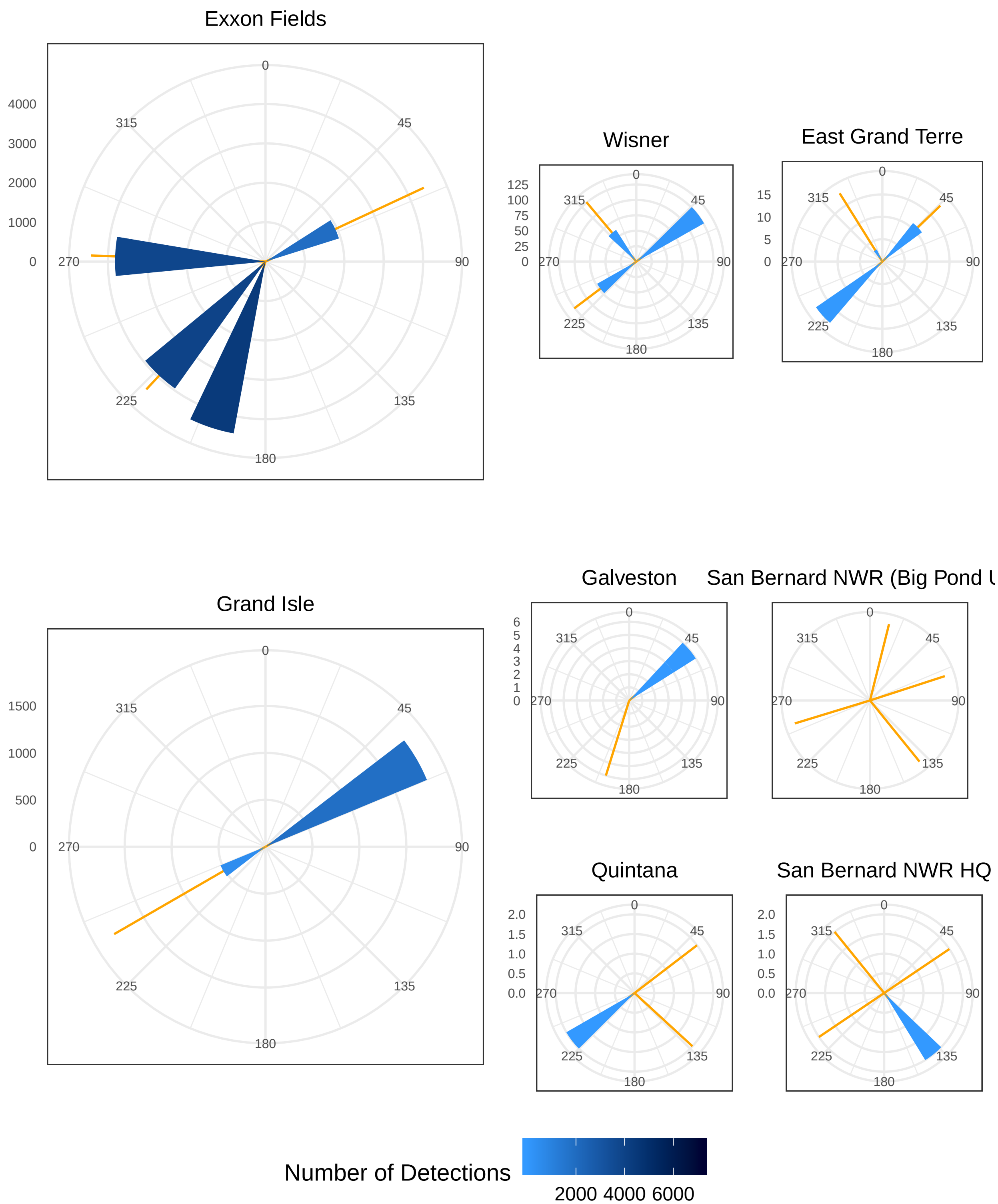


Figure 2.7: Frequency of all 2017 detections on each antenna. Increasingly darker blue bars indicate increased numbers of detections during the incubation period, orange lines indicate antenna bearings at each site. Note that scales differ by site.

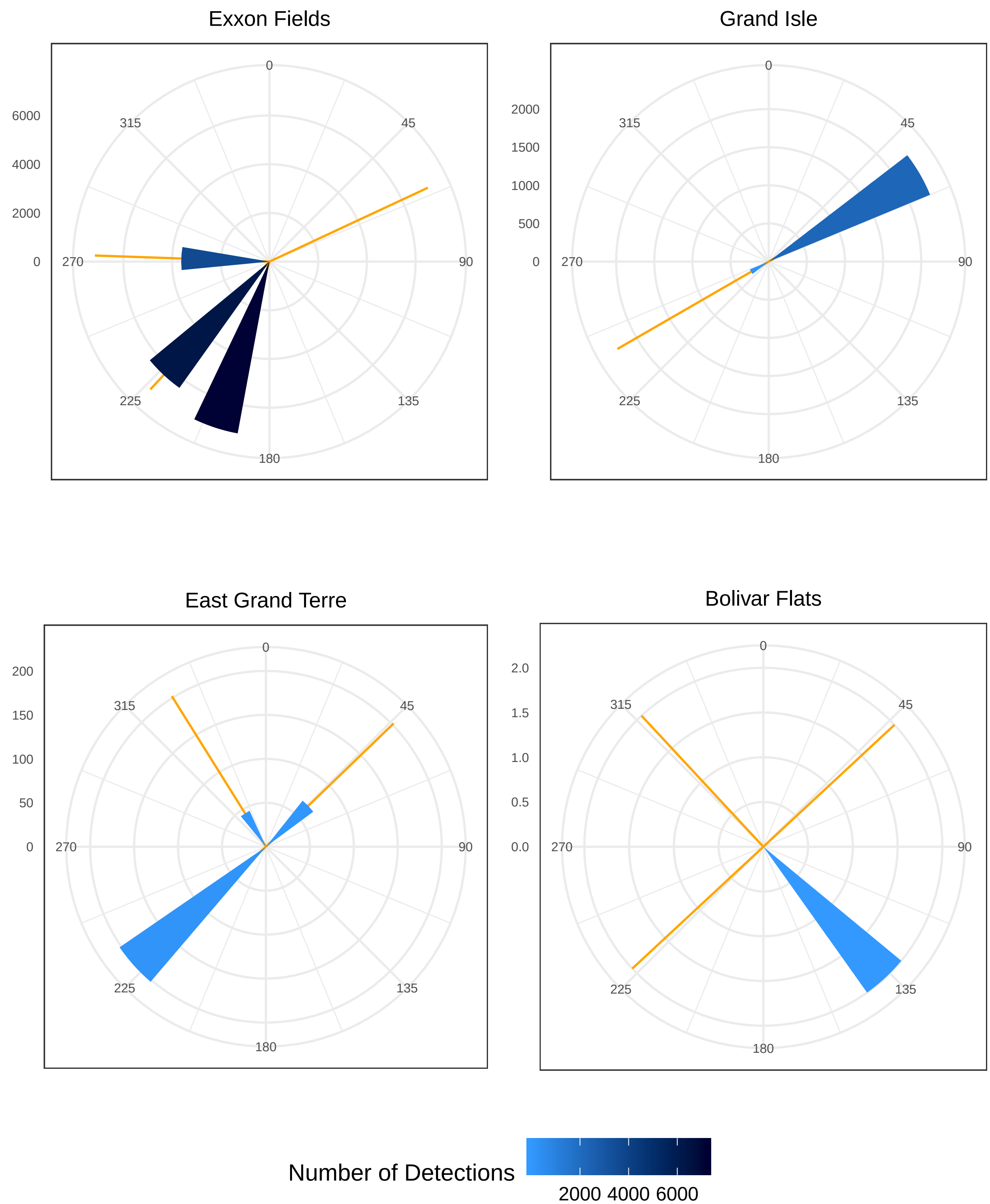


Figure 2.8: Frequency of all 2018 detections on each antenna. Increasingly darker blue bars indicate increased numbers of detections during the incubation period, orange lines indicate antenna bearings at each site. Note that scales differ by site.

Appendix A

Appendix - Downloading and Cleaning Data

Detailed instructions for accessing and understanding data can be found in the Motus R Book, available here.

A.1 Loading Packages

In order to download data and complete the analysis you will need a number of packages, if you do not already have them installed, follow the instructions below:

```
# Install devtools to allow download of Motus R
# packages
install.packages("devtools")

# Install motus for data download, data
# manipulation, visualization and analysis
install_github("MotusWTS/motus")

# install motusClient for data download
install_github("MotusWTS/motusClient")

# install packages for analysis
install.packages("ggmap")
install.packages("tidyr")
install.packages("gridExtra")
install.packages("ggplot2")
install.packages("viridis")
install.packages("knitr")
install.packages("pander")
```

Once packages are installed, load them into your project. To download and filter the data we will only need use of the motus R package which is loaded with the 'library' function:

```
# load motus package
library(motus)
```

A.2 Data Download - project 132

Downloading data from Motus requires you to enter your Motus username and password, you will only have access to data for projects to which you are registered. To register with Motus visit https://motus.org/.

To download data for the **first time** to your working directory:

```
proj.num <- 132  # assign the correct project number
sql.motus <- tagme(projRecv = proj.num, new = TRUE,
    update = TRUE)
```

To update an existing sql.motus file to your working directory:

```
proj.num <- 132  # assign the correct project number
sql.motus <- tagme(projRecv = proj.num, new = FALSE,
    update = TRUE)
```

A.3 Accessing Data

Data is downloaded as an SQL file, to analyze data the SQL file is converted into a flat file and only relevant columns and Least Tern tags are retained. At the same time a new column for receiver latitude and longitude is created, consisting of the GPS latitude if available, infilled with the deployment latitude if GPS is not available.

```
tbl.alltags <- tbl(sql.motus, "alltags") # convert sql file to a tbl
df.alltags <- tbl.alltags %>%
  mutate(recvLat = if_else((is.na(gpsLat)|gpsLat == 0), # create new lat/lon variables
                           recvDeployLat, gpsLat),
         recvLon = if_else((is.na(gpsLon)|gpsLon == 0),
                           recvDeployLon, gpsLon)) %>%
  #remove unnecessary columns
  select(-ambigID, -tagProjID, -sigsd, -noise, -freq, -freqsd, -slop, -burstSlop,
         -done, -bootnum, -tagType, -codeSet, -mfg, -tagModel, -tagLifespan,
         -nomFreq, -pulseLen, -markerNumber, -markerType, -tagDeployAlt,
         -tagDeployComments, -fullID, -recvDeployAlt, -antHeight, -speciesFR,
         -speciesSci, -speciesGroup, -tagProjName, gpsAlt, recvSiteName) %>%
  filter(speciesEN == "Least Tern") %>% # keep only LETE
  collect() %>%
  as.data.frame %>% # convert tbl to data.frame
  # convert times to datetime
  mutate(ts = as_datetime(ts, tz = "UTC", origin = "1970-01-01"),
         tagDeployStart = as_datetime(tagDeployStart, tz = "UTC", origin = "1970-01-01"),
         tagDeployEnd = as_datetime(tagDeployEnd, tz = "UTC", origin = "1970-01-01"))
```

A.4 Dataframe of motusTagID and Lotek tag ID

```
tagIDs <- unique(select(df.alltags, motusTagID, mfgID))
```

A.5 Cleaning Data

Projects with receivers in the tagging area are 31, 65, 120, and 132, initially all detections from receivers registered to these projects will be retained. For all other projects any detections with a runLen < 2 are deemed to be false positives. Filter 'blocks' are created each time a detection is identified as a false positive, at the end of the cleaning process all the false positive blocks are combined to create a single filter to apply to all the data.

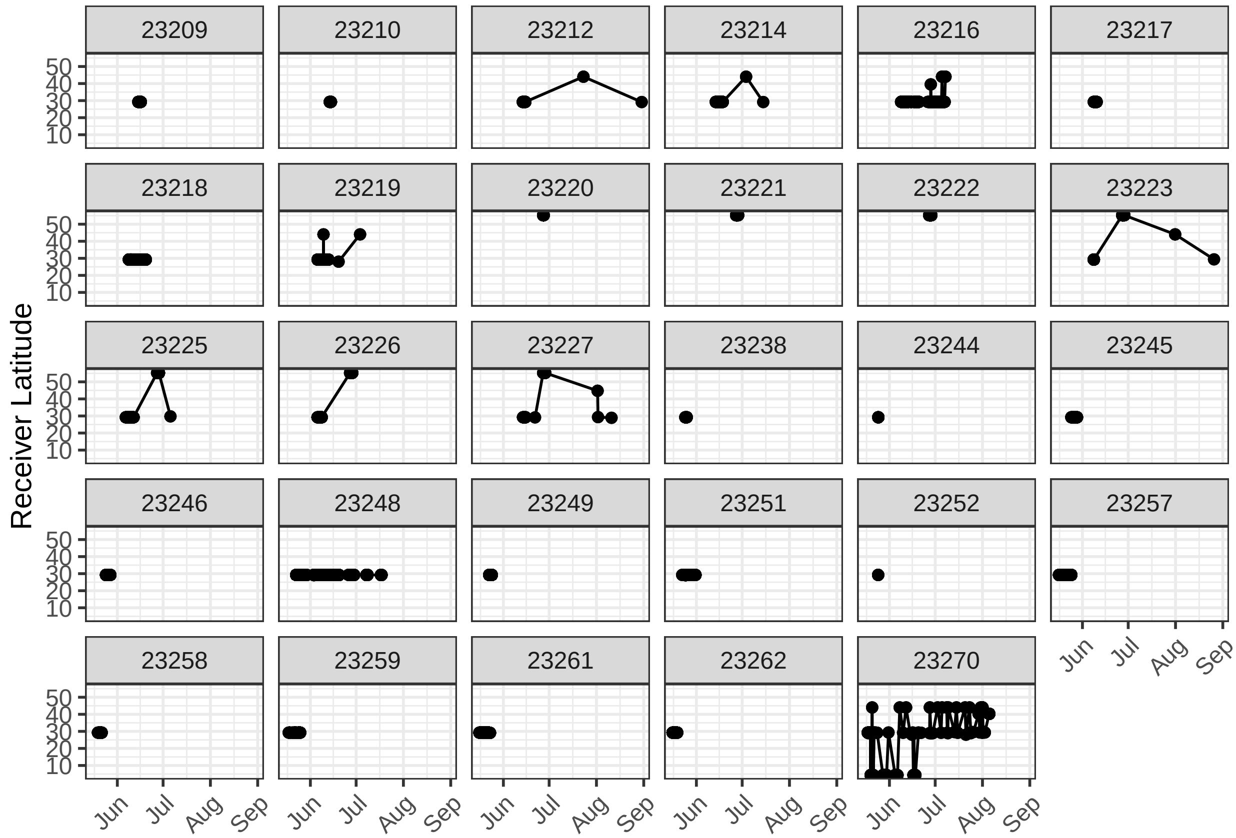
First, remove detections outside tagging area with a runLen of 2:

```
# create list of projects with receivers outside the tagging area
nonTaggingLocs <- unique(filter(df.alltags,
                              !(recvProjID %in% c(31, 65, 120, 132)))$recvProjID)
lete17 <- filter(df.alltags, year(ts) == 2017, # filter to only 2017 detections
                 !(recvProjID %in% nonTaggingLocs & runLen == 2)) %>%
# remove detections with a runLen <= 2 for stations outside tagging area
  arrange(motusTagID, ts)

# create the first filter block
df.block.0 <- filter(df.alltags, year(ts) == 2017,
                     recvProjID %in% nonTaggingLocs & runLen == 2)
```

Next, detections are plotted by latitude to visually identify other questionable detections to eventually create more blocks of false positives for the final filter. Below are detections for all 2017 tags by latitude:

```
# Create ggplot theme for plotting
th <- theme_bw() + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1), plot.title = element_text(hjust = 0.5))
# plot all tags by latitude
ggplot(lete17, aes(ts, recvLat)) + geom_point() + geom_path() +
    facet_wrap(~motusTagID) + th + labs(x = NULL, y = "Receiver Latitude")
```

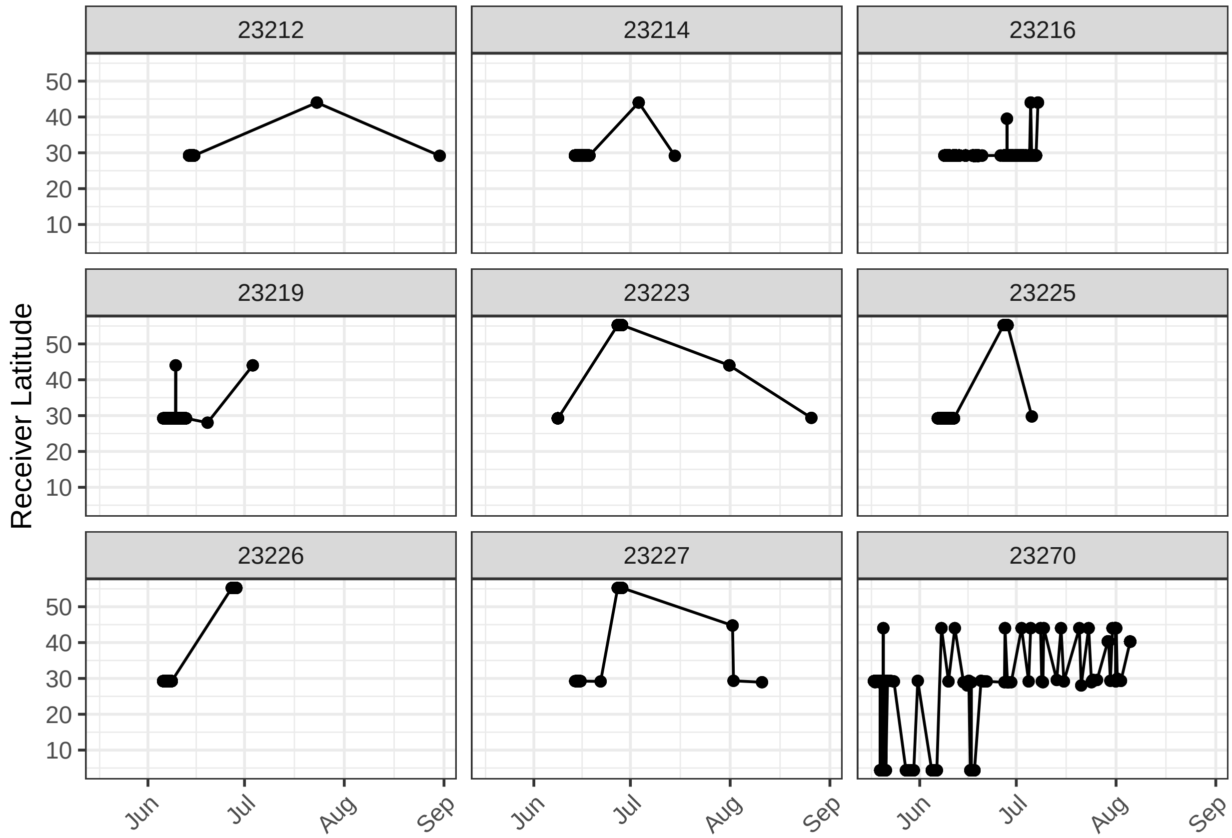


Tags are removed that were never detected at the breeding site, these are **23220**, **23221**, and **23222**.

```
# remove tags 23220, 23221, 23222 and create a
# block
lete17 <- filter(lete17, !(motusTagID %in% c(23220,
    23221, 23222)))
df.block.1 <- filter(df.alltags, year(ts) == 2017 &
    motusTagID %in% c(23220, 23221, 23222))
```

Next, tags that were detected outside the breeding area are examined: **23212, 23214, 23216, 23219, 23223, 23225, 23226, 23227, 23270**. Below is a plot of these tags, and a table displaying the number of runs with runLen < 5 for each station:

```
ggplot(filter(lete17, motusTagID %in% c(23212, 23214,
    23216, 23219, 23223, 23225, 23226, 23227, 23270)),
    aes(ts, recvLat)) + geom_point() + geom_path() +
    facet_wrap(~motusTagID) + th + labs(x = NULL, y = "Receiver Latitude")
```




```
# examining runLen < 5
filter(lete17, motusTagID %in% c(23212, 23214, 23216,
    23219, 23223, 23225, 23226, 23227, 23270), runLen <
    5) %>% group_by(recvDeployName, runLen) %>% summarize(n = length(motusTagID)) %>%
    as.data.frame() %>% arrange(runLen, n)
```

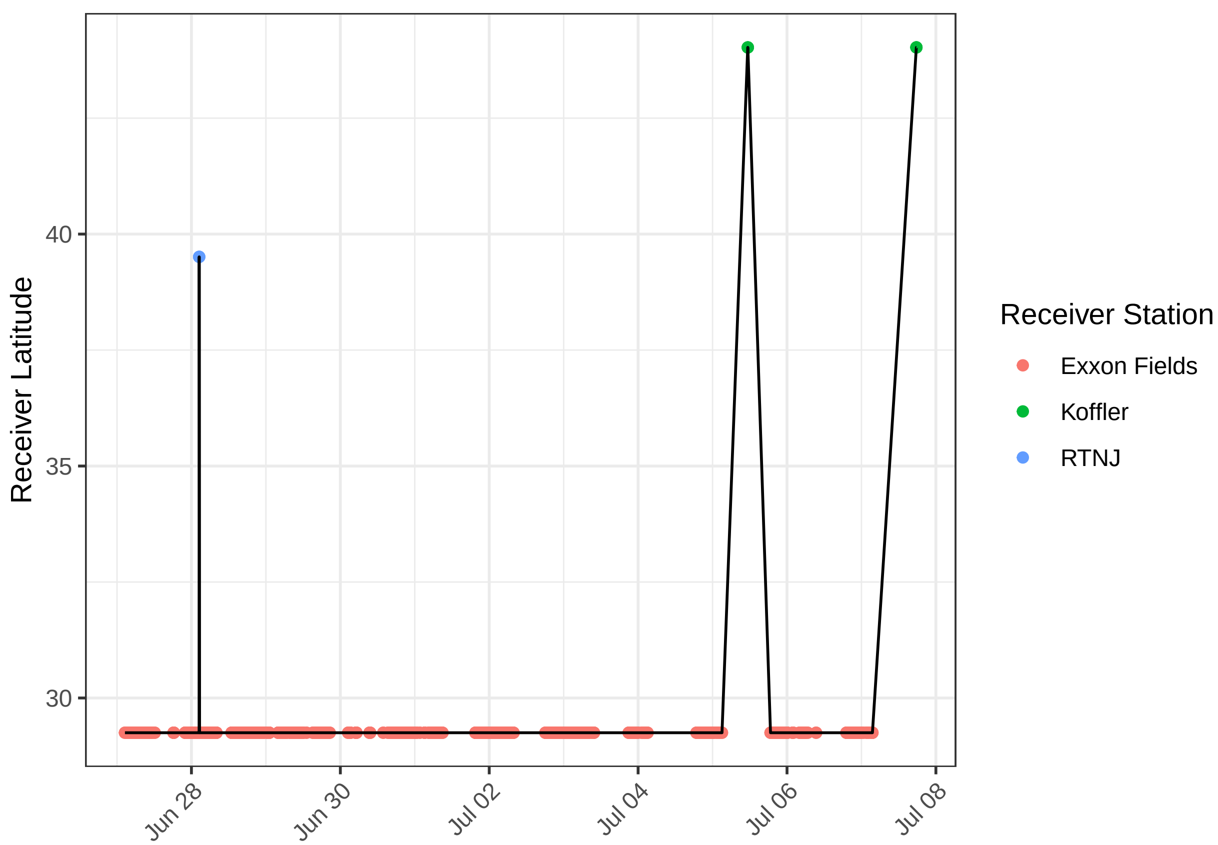
```
##                                    recvDeployName runLen    n
## 1                                    Bolivar_Flats      2    2
## 2                                     Peveto Woods      2    2
## 3  San Bernard National Wildlife Refuge Headquarters      2    2
## 4                                       Sabine NWR      2    4
## 5                    Rockport Bay Education Center      2    6
## 6                                      High_Island      2    8
## 7                                           Wisner      2    8
## 8                                 Scenic Galveston      2   24
## 9                                         Quintana      2   28
## 10                    San Bernard NWR (Big Pond Unit)      2   30
## 11                                      Grand_Isle      2   82
## 12                                    Exxon Fields      2 4688
## 13                                            RTNJ      3    3
## 14                                      Swallowtail      3    3
## 15                                         Drasher      3    9
## 16                                          Wisner      3   12
## 17                                    Waggoners Gap      3   60
## 18                                         Koffler      3   72
## 19                                      Grand_Isle      3  120
```

```
## 20                                       Burntpoint      3   258
## 21                               Los Vientos Forest      3   411
## 22                                     Exxon Fields      3  5736
## 23                                          Koffler      4    12
## 24                                     Waggoners Gap     4    12
## 25                               Los Vientos Forest      4    32
## 26                                       Grand_Isle      4    72
## 27                                     Exxon Fields      4  4952
```

Stations that stand out as having a large number of runLen = 3, far outside the tagging area include: *Koffler*, *Waggoners Gap*, *Burntpoint*, and *Los Vientos Forest*. These are likely sites with a large amount of background radio noise leading to increased levels of false positives. To confirm, a few examples are provided below.

For **tag 23216**, given the amount of time and distance between detections at *RTNJ* and *Exxon Fields*, and between *Koffler* and *Exxon Fields*, these are not biologically possible flights and these detections will be removed:

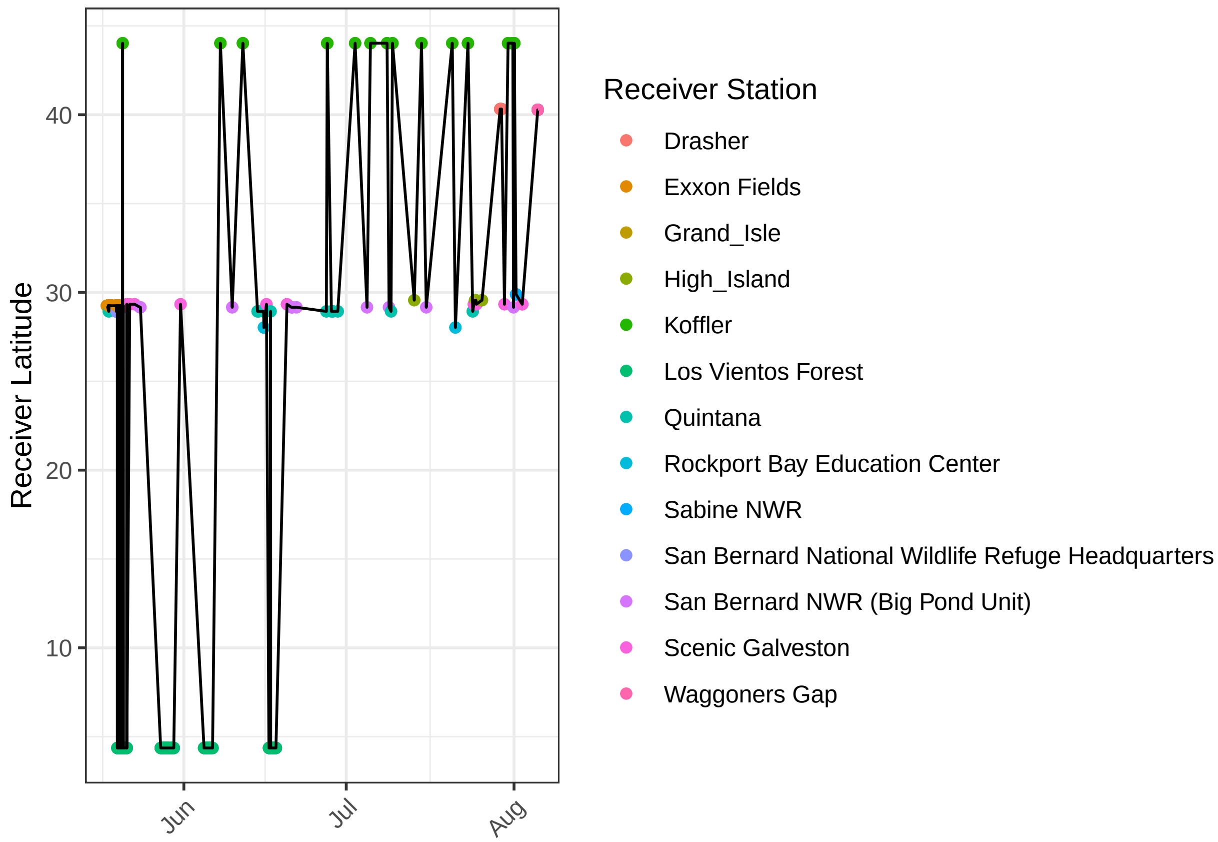
```
# motusTagID 23216
ggplot(filter(lete17, motusTagID == 23216, ts > as.POSIXct("2017-06-26")),
    aes(ts, recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



For **tag 23270**, it again appears there are clear false positives that can be removed at *Koffler*, *Los Vientos Forest*, *Drasher*, and *Waggoners Gap*. These detections will therefore be flagged as false positives:

```
# motusTagID 23270
ggplot(filter(lete17, motusTagID == 23270), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
```

```
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```

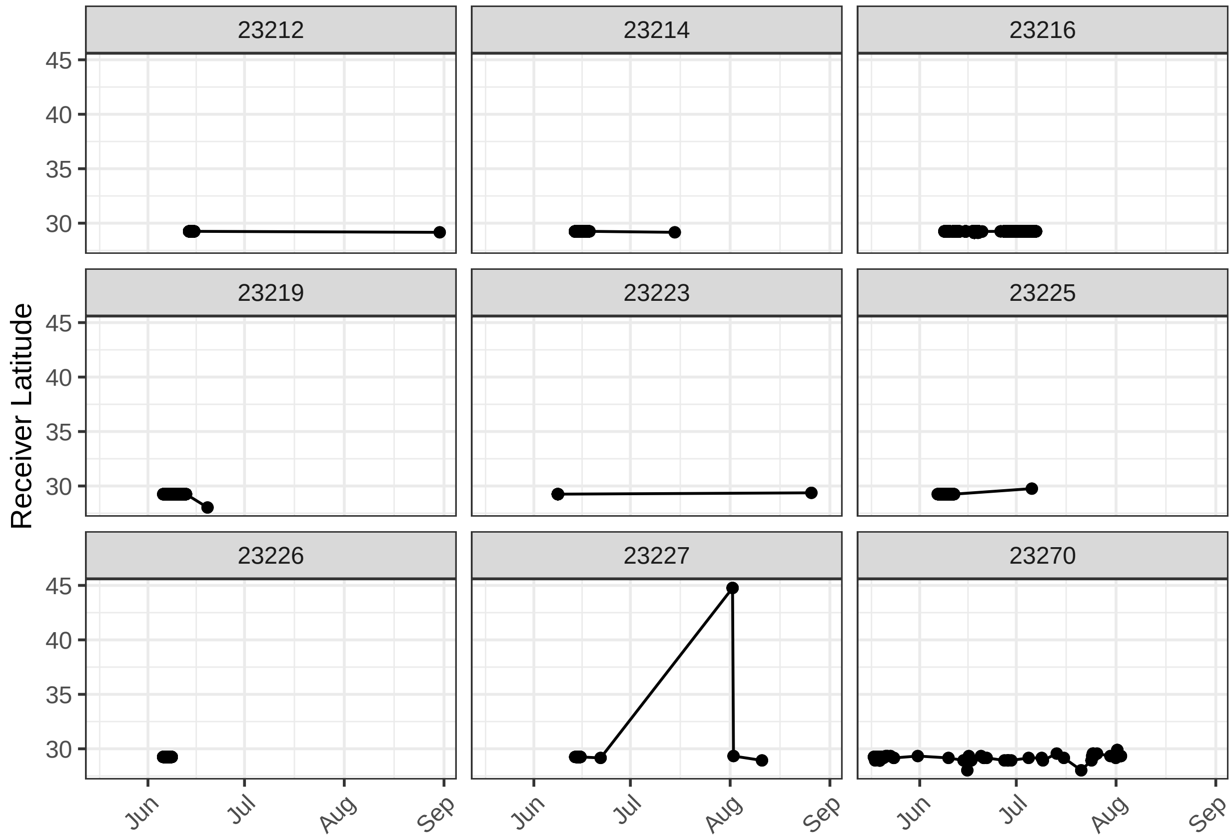


Create a filter block to remove noisy sites with high amounts of short runLen's:

```
df.block.2 <- filter(df.alltags, year(ts) == 2017,
    recvDeployName %in% c("Drasher", "RTNJ", "Koffler",
        "Waggoners Gap", "Burntpoint", "Los Vientos Forest"))
```

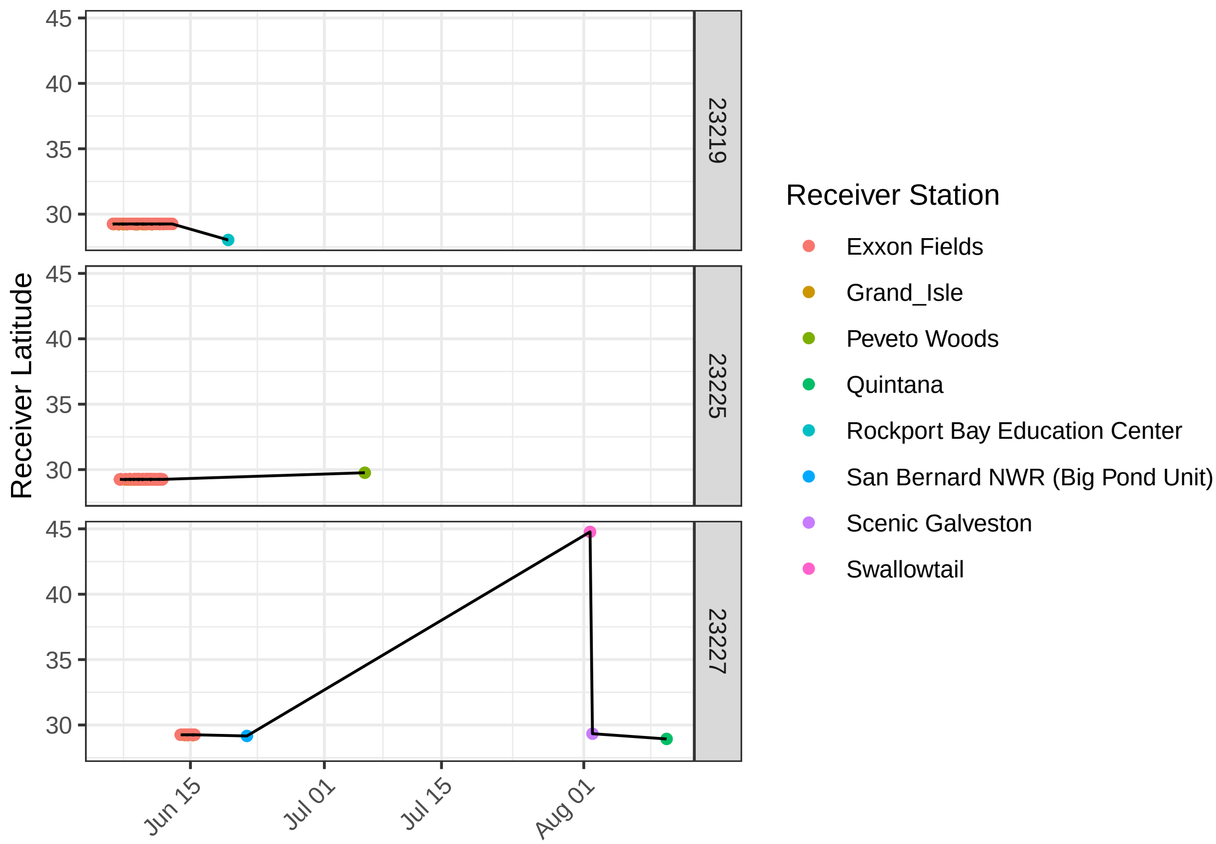
Removing the sites with a high number of low runLen's results in much cleaner plots:

```
lete17 <- filter(lete17, !(recvDeployName %in% c("Drasher",
    "RTNJ", "Koffler", "Waggoners Gap", "Burntpoint",
    "Los Vientos Forest")))
ggplot(filter(lete17, motusTagID %in% c(23212, 23214,
    23216, 23219, 23223, 23225, 23226, 23227, 23270)),
    aes(ts, recvLat)) + geom_point() + geom_path() +
    facet_wrap(~motusTagID) + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



There are three remaining tags with questionable detections: **23219, 23225, 23227**:

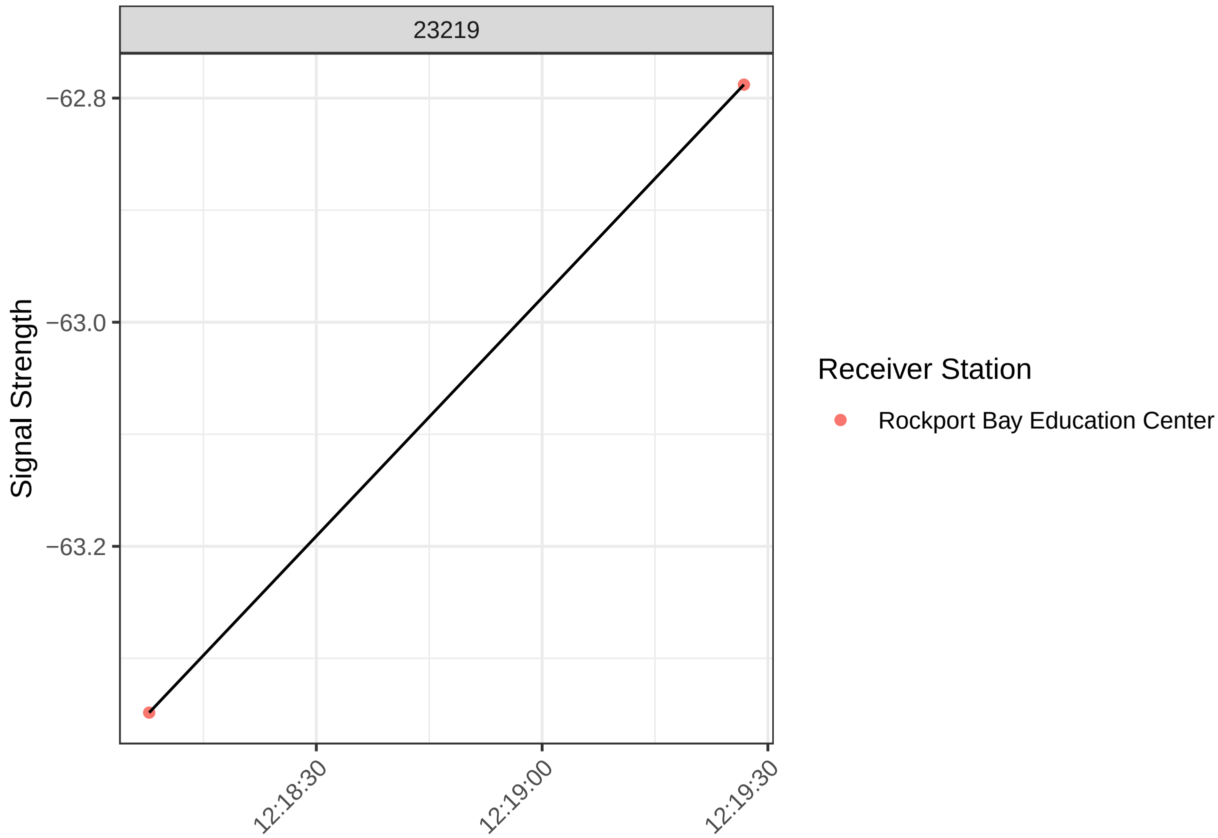
```
ggplot(filter(lete17, motusTagID %in% c(23219, 23225,
    23227)), aes(ts, recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + facet_grid(motusTagID ~ .) + th +
    labs(x = NULL, y = "Receiver Latitude", col = "Receiver Station")
```



To determine if detections are reliable, each tag will be examined on an individual basis.

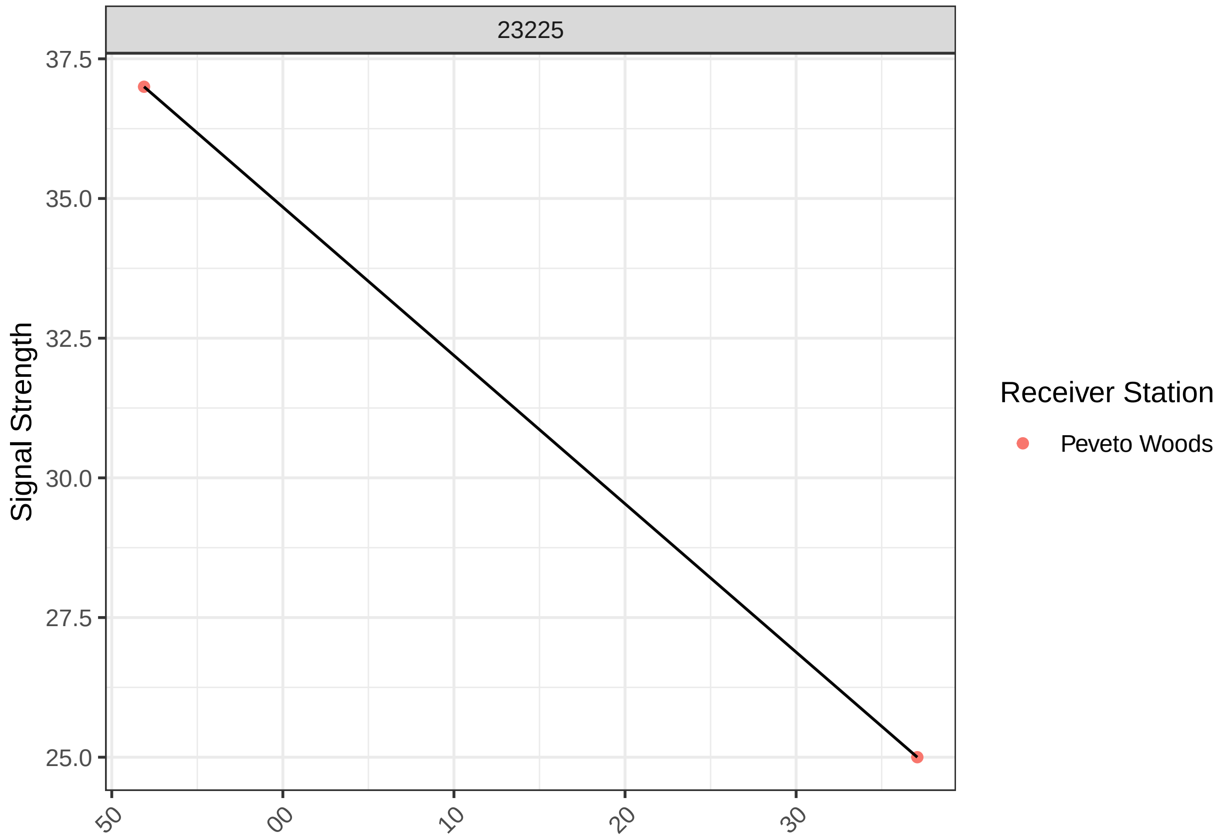
Tag 23219: Detections at Rockport Bay Education Center are only runLen 2, these are not reliable detections:

```
ggplot(filter(lete17, motusTagID == 23219, ts > as.POSIXct("2017-06-18")),
    aes(ts, sig)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + facet_wrap(~motusTagID) + th + labs(x = NULL,
    y = "Signal Strength", col = "Receiver Station")
```



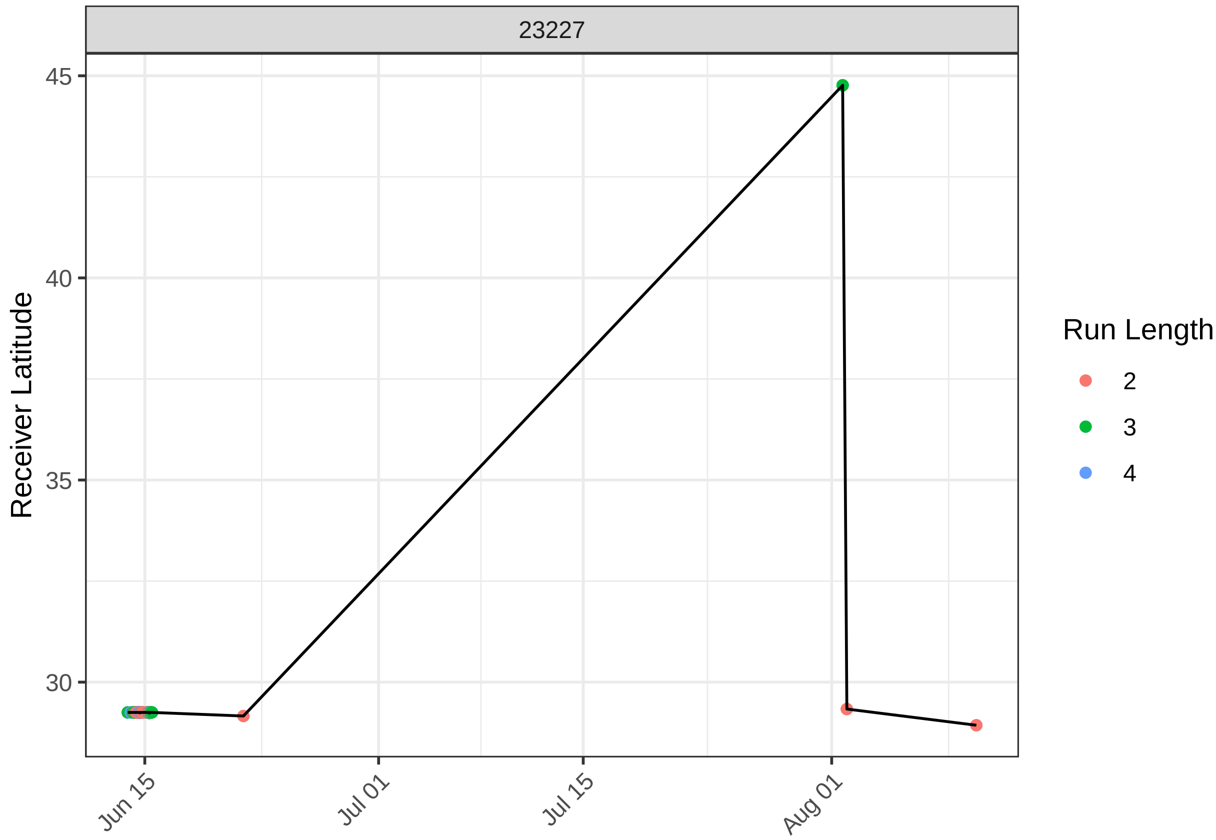
Tag 23225: detections at Peveto Woods are only runLen 2, these are not reliable detections:

```
ggplot(filter(lete17, motusTagID == 23225, ts > as.POSIXct("2017-07-02")),
    aes(ts, sig)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + facet_wrap(~motusTagID) + th + labs(x = NULL,
    y = "Signal Strength", col = "Receiver Station")
```



Tag 23227: all detections after June 15 are of runLen 2, except for detections at Swallowtail (receiver latitude = 45, runLen = 3) which are too far away to be legitimate detections so these will be removed.

```
ggplot(filter(lete17, motusTagID == 23227), aes(ts,
    recvLat)) + geom_point(aes(col = as.factor(runLen))) +
    geom_path() + facet_wrap(~motusTagID) + th + labs(x = NULL,
    y = "Receiver Latitude", col = "Run Length")
```



A filter block is created to remove false detections for the three tags above:

```
df.block.3 <- filter(df.alltags, year(ts) == 2017,
    motusTagID == 23219 & recvDeployName == "Rockport Bay Education Center")
df.block.4 <- filter(df.alltags, year(ts) == 2017,
    motusTagID == 23225 & recvDeployName == "Peveto Woods")
df.block.5 <- filter(df.alltags, year(ts) == 2017,
    motusTagID == 23227 & ts > as.POSIXct("2017-06-15"))
```

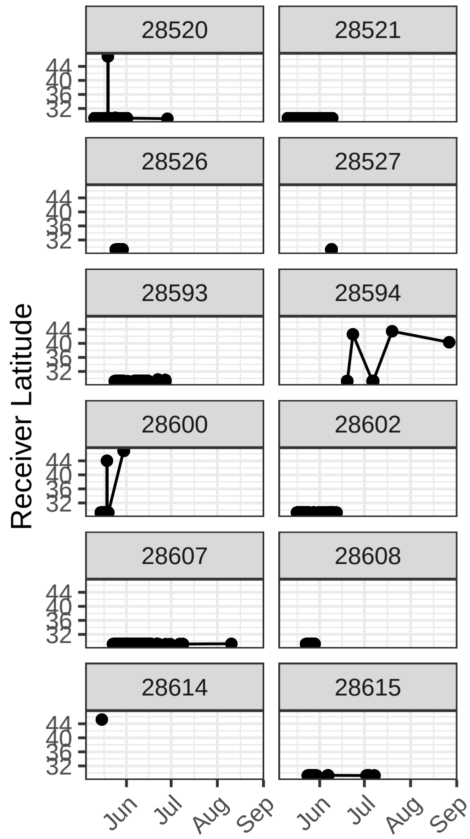
The process is repeated with 2018 detection data, starting by removing detections from stations outside the tagging area with a runLen < 2:

```
# create list of projects with receivers outside the tagging area
lete18 <- filter(df.alltags, year(ts) == 2018, # filter to only 2018 detections
                 !(recvProjID %in% nonTaggingLocs & runLen == 2)) %>%
# remove detections with a runLen < 2 for stations outside tagging area
  arrange(motusTagID, ts)
# create the first filter block
df.block.6 <- filter(df.alltags, year(ts) == 2018,
                     recvProjID %in% nonTaggingLocs & runLen == 2)
```

Next, all tags are plotted by latitude to visually identify potential false positives.

Plot of all 2018 tags by latitude:

```
ggplot(lete18, aes(ts, recvLat)) + geom_point() + geom_path() +
    facet_wrap(~motusTagID) + th + labs(x = NULL, y = "Receiver Latitude")
```

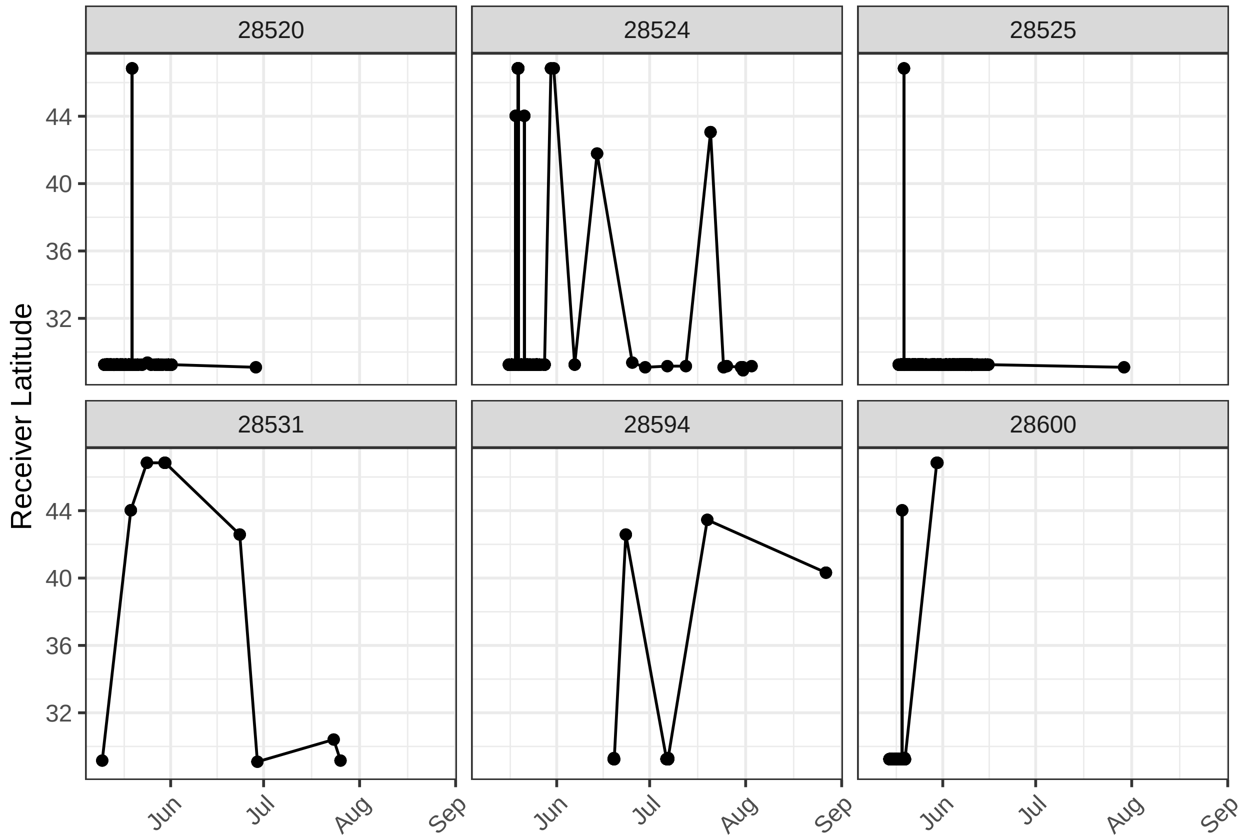


Tags that were never detected at the breeding site are removed and a filter block is created: **28614**

```
lete18 <- filter(lete18, !(motusTagID == 28614))
df.block.7 <- filter(df.alltags, year(ts) == 2018,
    motusTagID == 28614)
```

Next tags that were detected outside breeding area were examined: **28520, 28524, 28525, 28531, 28594, 28600**

```
ggplot(filter(lete18, motusTagID %in% c(28520, 28524,
    28525, 28531, 28594, 28600)), aes(ts, recvLat)) +
    geom_point() + geom_path() + facet_wrap(~motusTagID) +
    th + labs(x = NULL, y = "Receiver Latitude")
```



Again, runLen for these detections is examined:

```
filter(lete18, motusTagID %in% c(28520, 28524, 28525,
    28531, 28594, 28600), runLen < 5) %>% group_by(recvDeployName,
    runLen) %>% summarize(n = length(motusTagID)) %>%
    as.data.frame() %>% arrange(runLen, n)
```

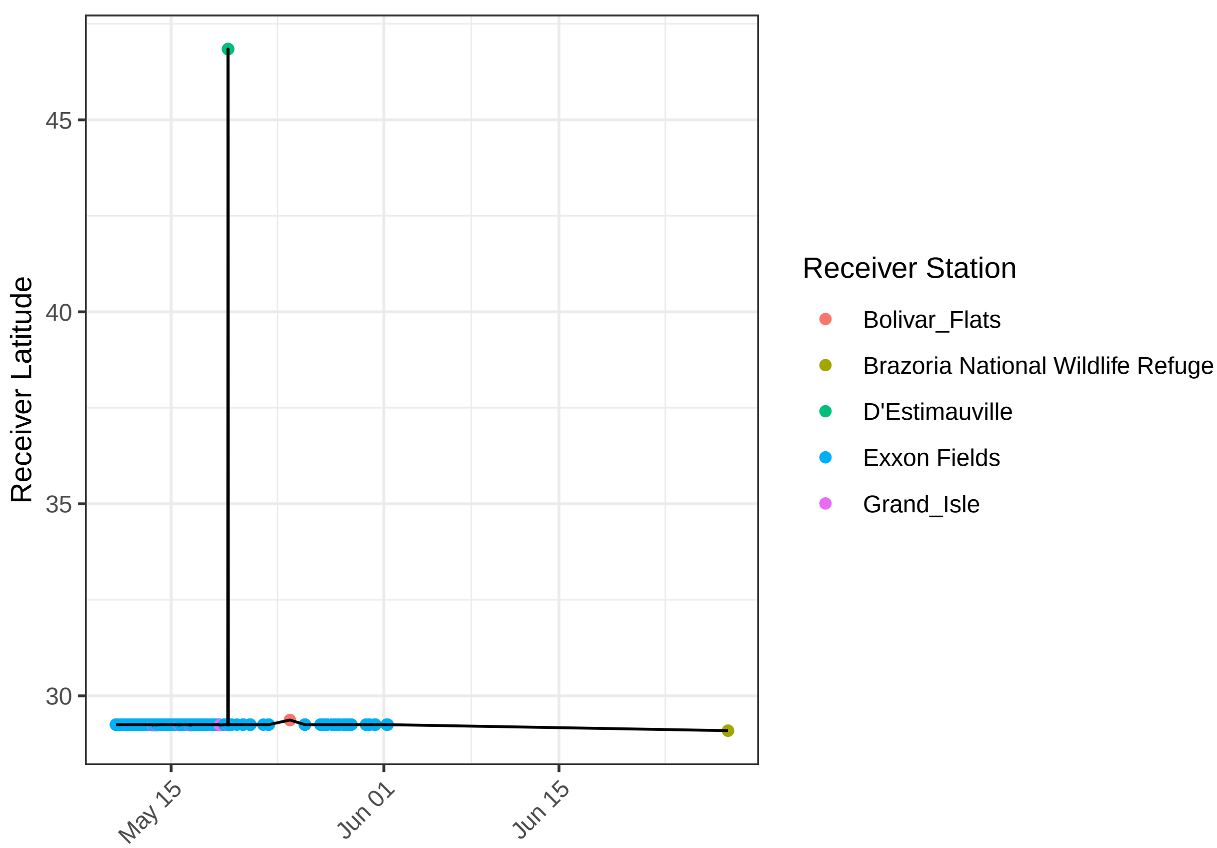
```
##                                       recvDeployName runLen    n
## 1  San Bernard National Wildlife Refuge Headquarters      2    2
## 2                        Trinity River NWR Die House      2    2
## 3                                      Bolivar_Flats      2    4
## 4                                   East Grand Terre      2    6
## 5                       San Bernard NWR (Big Pond Unit)      2   14
## 6                    Brazoria National Wildlife Refuge      2   18
## 7                                         Grand_Isle      2   82
## 8                                       Exxon Fields      2 3174
## 9                                               BPLH      3    3
## 10                                           Curries      3    3
## 11                                           Drasher      3    3
## 12                                        Stutchbury      3    3
## 13                                           Old Cut      3    6
## 14                                           Koffler      3   12
## 15                                        Grand_Isle      3   48
## 16                                    D'Estimauville      3  159
## 17                                      Exxon Fields      3 2391
## 18                                           Koffler      4    8
## 19                                    D'Estimauville      4   20
```

```
## 20                                     Grand_Isle      4   32
## 21                                   Exxon Fields      4 1592
```

Stations that stand out as having a large number of runLen = 3, far outside the tagging area include:*D'Estimauville*, and as in 2017 *Koffler*. These are likely sites with a large amount of background radio noise leading to increased levels of false positives. Tags detected at these stations are examined individually to remove false positives.

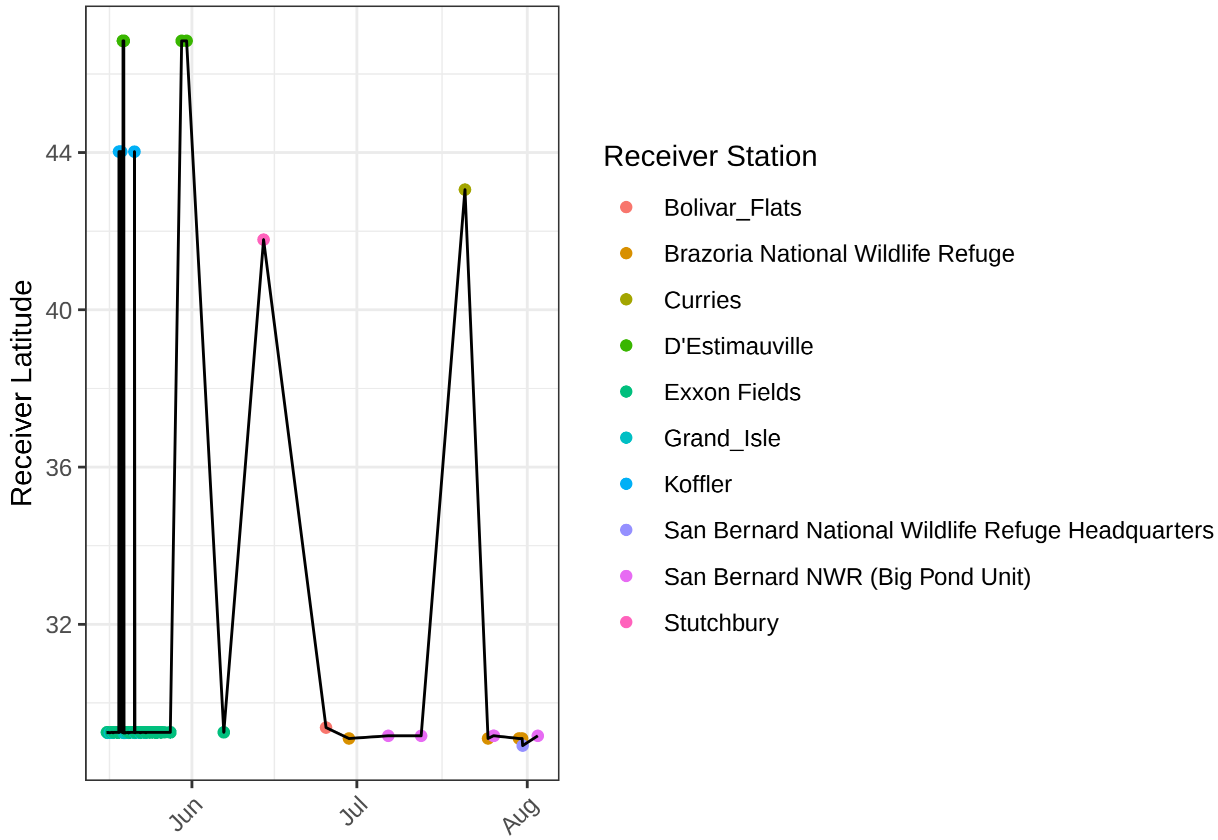
Tag 28520: The detections at *D'Estimauville* in Quebec are not biologically possible given the timing, and was also flagged as a noisy site, these detections are flagged as false positives.

```
ggplot(filter(lete18, motusTagID == 28520), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



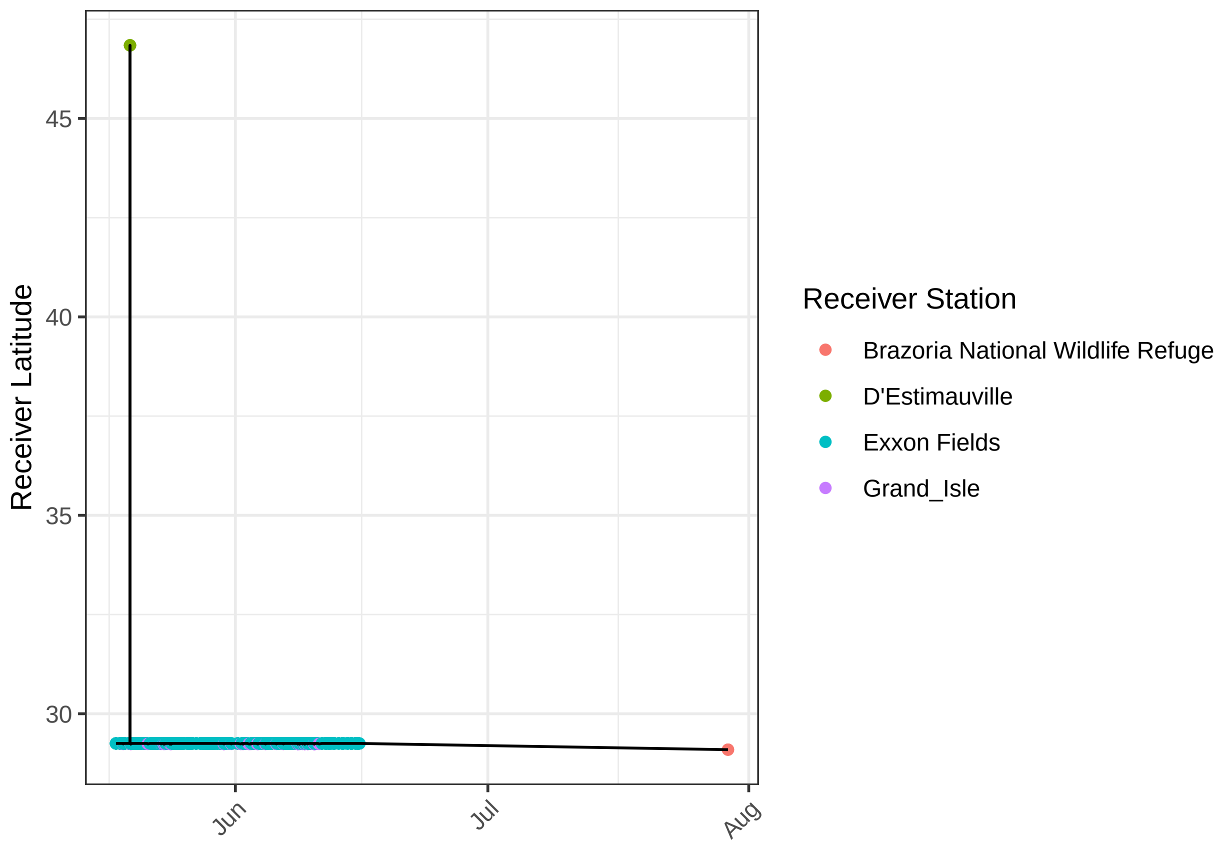
Tag 28524: Given the amount of time and distance between detections at *Koffler* and *Exxon Fields*, and between *D'Estimauville* and *Exxon Fields*, these are not biologically possible flights, they were also flagged as noisy stations. The detections at Stutchbury and Curries are of runLen 3, and occur in Ontario and Pennsylvania respectively. These detections are therefore flagged as false positives.

```
ggplot(filter(lete18, motusTagID == 28524), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



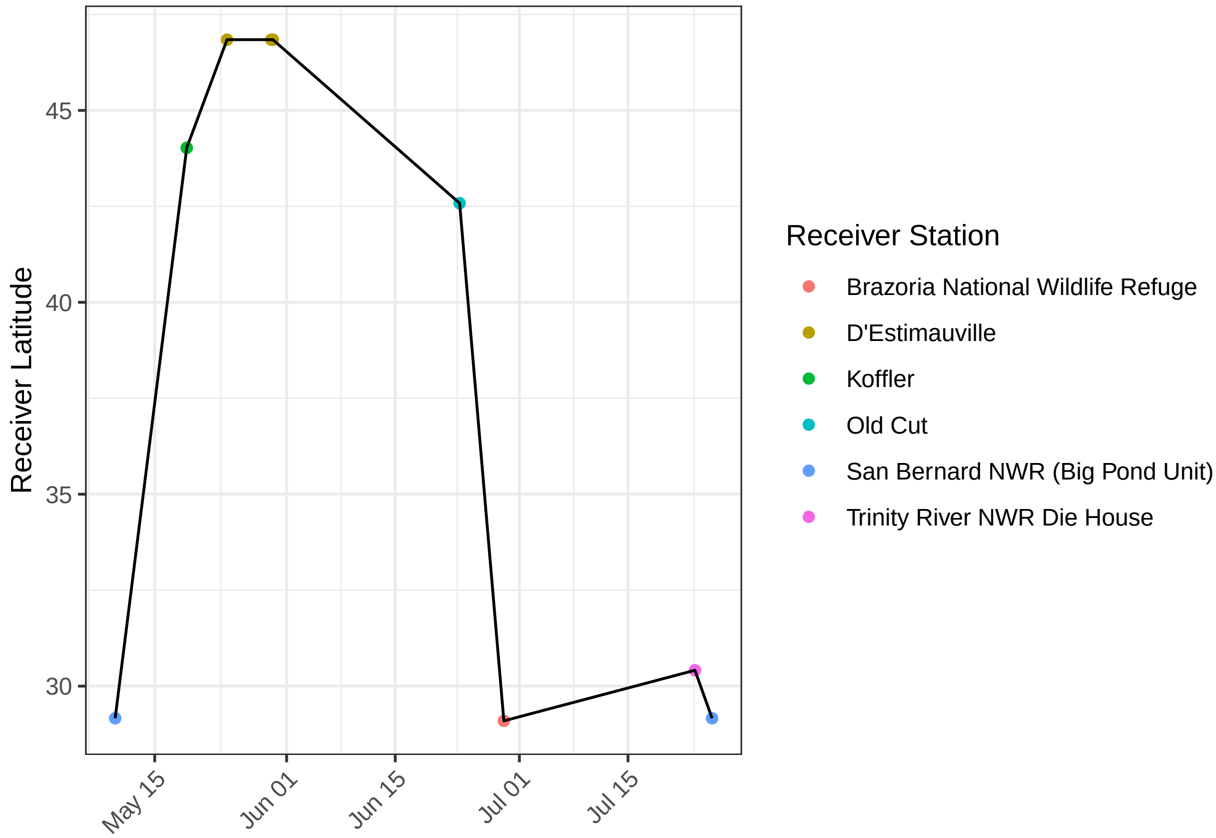
Tag 28525: Just as with tag 28520, the detection at *D'Estimauville* will be removed.

```
ggplot(filter(lete18, motusTagID == 28525), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



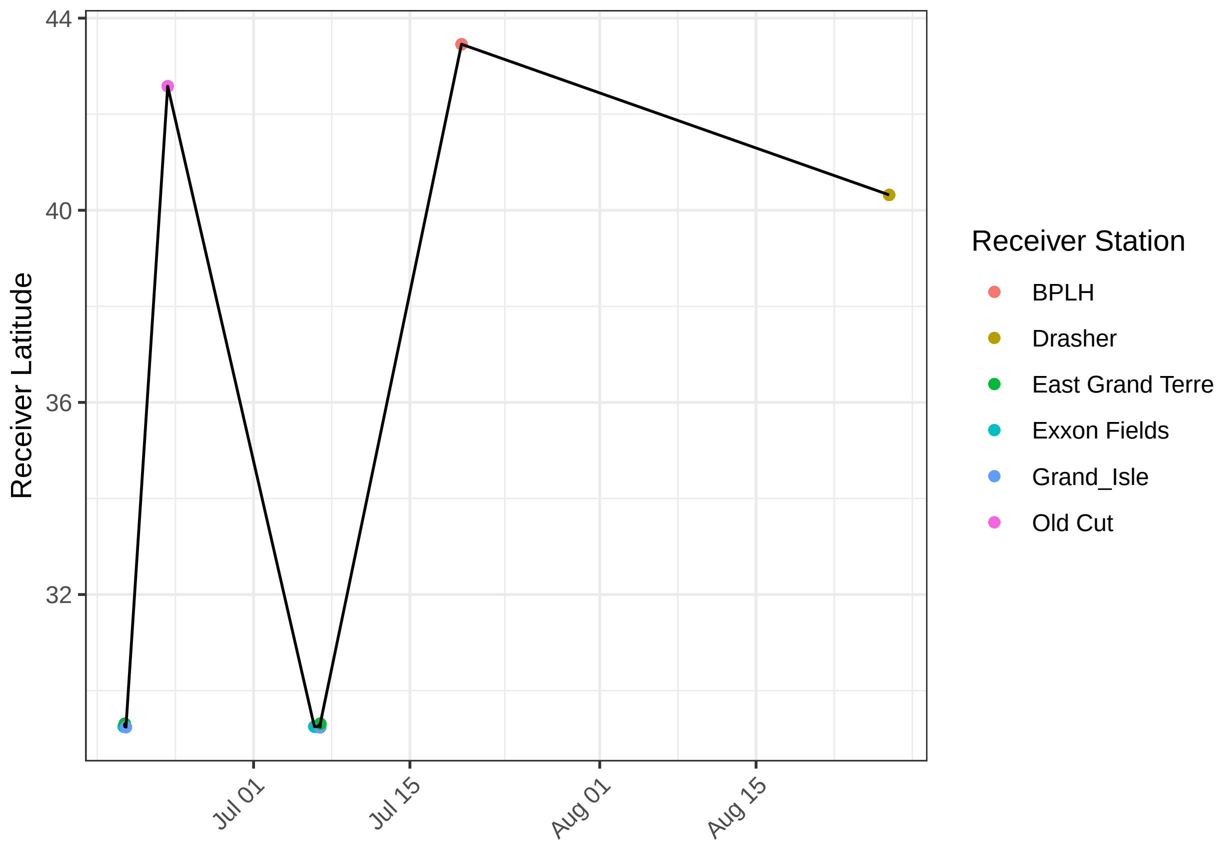
Tag 28531: Again, it is unlikely that this tag was in Ontario and Quebec in June so detections at *Koffler*, *Old Cut*, and *D'Estimauville* will be removed.

```
ggplot(filter(lete18, motusTagID == 28531), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



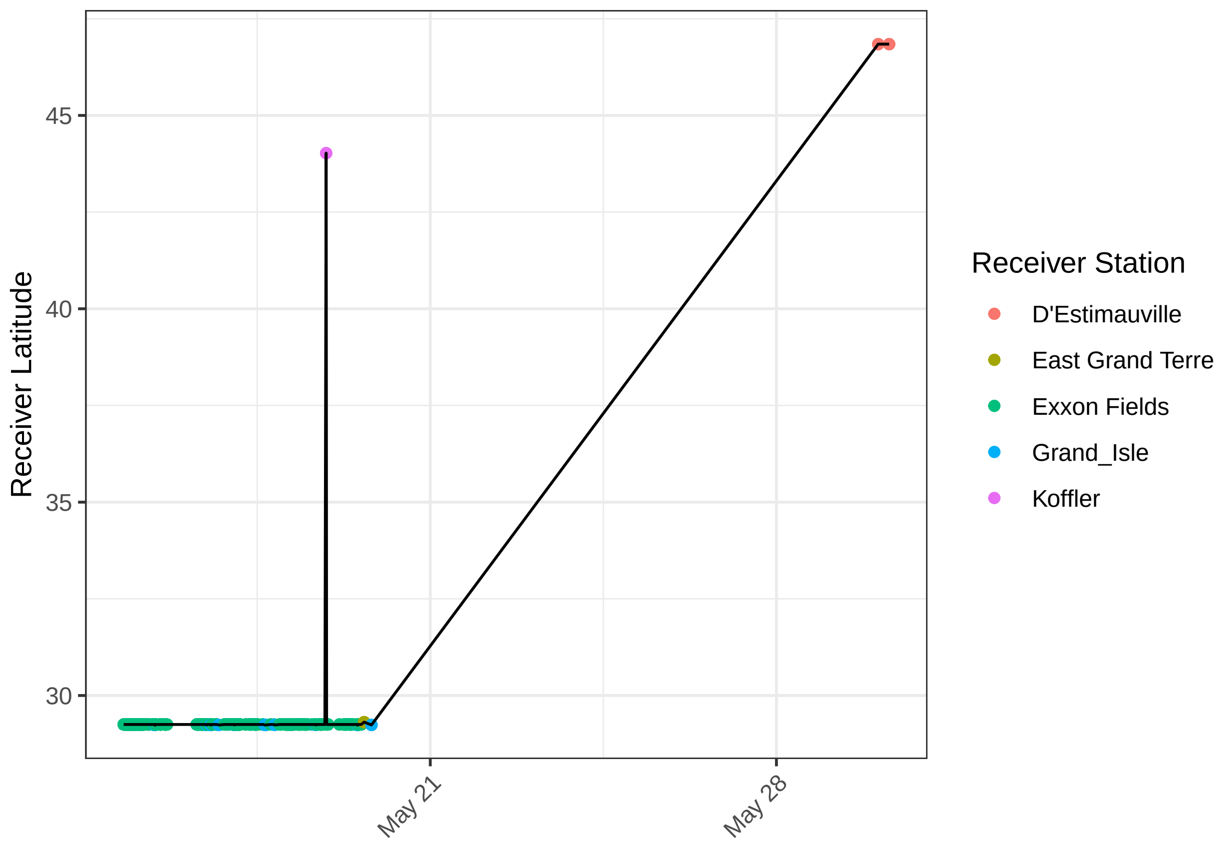
Tag 28594: Again, it is unlikely that this tag was in Nova Scotia, Ontario, or Pennsylvania (stations *BPLH*, *Old Cut*, and *Drasher* respectively), and all runLen are 3, so detections at these stations will be removed.

```
ggplot(filter(lete18, motusTagID == 28594), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```



Tag 28600: Detections at *Koffler* and *D'Estimauville* once again appear problematic and will be removed.

```
ggplot(filter(lete18, motusTagID == 28600), aes(ts,
    recvLat)) + geom_point(aes(col = recvDeployName)) +
    geom_path() + th + labs(x = NULL, y = "Receiver Latitude",
    col = "Receiver Station")
```

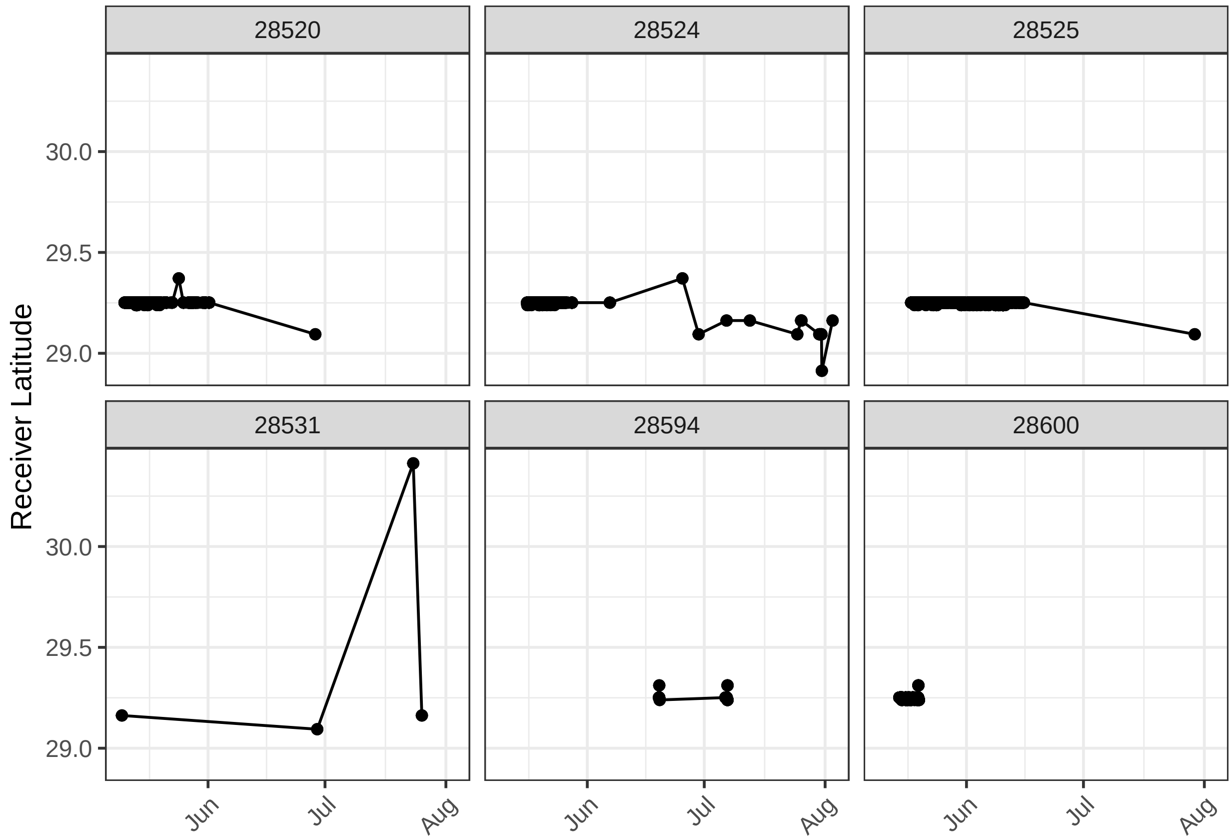


Since all false positives in these cases are for the same stations, one filter block is used to apply to these false detections:

```
df.block.8 <- filter(df.alltags, year(ts) == 2018,
    recvDeployName %in% c("D'Estimauville", "Koffler",
        "Stutchbury", "Curries", "Old Cut", "BPLH",
        "Drasher"))
```

Here the plots are redone after removing the above identified false positives:

```
lete18 <- filter(lete18, !(recvDeployName %in% c("D'Estimauville",
    "Koffler", "Stutchbury", "Curries", "Old Cut",
    "BPLH", "Drasher")))
ggplot(filter(lete18, motusTagID %in% c(28520, 28524,
    28525, 28531, 28594, 28600)), aes(ts, recvLat)) +
    geom_point() + geom_path() + facet_wrap(~motusTagID) +
    th + labs(x = NULL, y = "Receiver Latitude")
```



Now that filter blocks have been created for all false positives, a single custom filter is created for all false positives. The Motus R package offers functionalities to save your filters directly within your .motus file. Once they're saved in your database, you can do the type of left_join as below without having to rely on data frames or an RDS file to store your data. More info can be found in Appendix D of the Motus R Book.

```
# Combine our df.blocks into a single data frame,
# and apply a probability of 0 for filtered
# records.
df.block.all <- bind_rows(df.block.0, df.block.1, df.block.2,
    df.block.3, df.block.4, df.block.5, df.block.6,
    df.block.7, df.block.8) %>% mutate(probability = 0)

# Create a new filter with the name filtFalsePos
# and populate it with df.block.all
tbl.filter = writeRunsFilter(sql.motus, "filtFalsePos",
    df = df.block.all, delete = TRUE)
```

```
##
## Filter records saved
```

```
# Obtain a table object where the filtered records
# from tbl.filter have been removed
tbl.alltags.sub <- left_join(tbl.alltags, tbl.filter,
    by = c("runID", "motusTagID")) %>% mutate(probability = ifelse(is.na(probability),
    1, probability)) %>% filter(probability > 0)
```

Appendix B

Appendix - Distances Between Receivers

The following table lists the latitude and longitude for each pair of receivers (X and Y), and distance (miles) between them:

Table B.1: Location and distance between each pair of receivers

Receiver X	Receiver Y	Location X
Bolivar_Flats	Bolivar_Flats	29.37, -94.75
Bolivar_Flats	Brazoria National Wildlife Refuge	29.37, -94.75
Bolivar_Flats	East Grand Terre	29.37, -94.75
Bolivar_Flats	Exxon Fields	29.37, -94.75
Bolivar_Flats	Grand_Isle	29.37, -94.75
Bolivar_Flats	High_Island	29.37, -94.75
Bolivar_Flats	Quintana	29.37, -94.75
Bolivar_Flats	Rock W	29.37, -94.75
Bolivar_Flats	Rockport Bay Education Center	29.37, -94.75
Bolivar_Flats	Sabine NWR	29.37, -94.75
Bolivar_Flats	San Bernard National Wildlife Refuge Headquarters	29.37, -94.75
Bolivar_Flats	San Bernard NWR (Big Pond Unit)	29.37, -94.75
Bolivar_Flats	Scenic Galveston	29.37, -94.75
Bolivar_Flats	Wisner	29.37, -94.75
Brazoria National Wildlife Refuge	Bolivar_Flats	29.09, -95.28
Brazoria National Wildlife Refuge	Brazoria National Wildlife Refuge	29.09, -95.28
Brazoria National Wildlife Refuge	East Grand Terre	29.09, -95.28
Brazoria National Wildlife Refuge	Exxon Fields	29.09, -95.28
Brazoria National Wildlife Refuge	Grand_Isle	29.09, -95.28
Brazoria National Wildlife Refuge	High_Island	29.09, -95.28
Brazoria National Wildlife Refuge	Quintana	29.09, -95.28
Brazoria National Wildlife Refuge	Rock W	29.09, -95.28
Brazoria National Wildlife Refuge	Rockport Bay Education Center	29.09, -95.28
Brazoria National Wildlife Refuge	Sabine NWR	29.09, -95.28
Brazoria National Wildlife Refuge	San Bernard National Wildlife Refuge Headquarters	29.09, -95.28
Brazoria National Wildlife Refuge	San Bernard NWR (Big Pond Unit)	29.09, -95.28
Brazoria National Wildlife Refuge	Scenic Galveston	29.09, -95.28
Brazoria National Wildlife Refuge	Wisner	29.09, -95.28
East Grand Terre	Bolivar_Flats	29.31, -89.86
East Grand Terre	Brazoria National Wildlife Refuge	29.31, -89.86
East Grand Terre	East Grand Terre	29.31, -89.86
East Grand Terre	Exxon Fields	29.31, -89.86
East Grand Terre	Grand_Isle	29.31, -89.86
East Grand Terre	High_Island	29.31, -89.86
East Grand Terre	Quintana	29.31, -89.86
East Grand Terre	Rock W	29.31, -89.86
East Grand Terre	Rockport Bay Education Center	29.31, -89.86
East Grand Terre	Sabine NWR	29.31, -89.86
East Grand Terre	San Bernard National Wildlife Refuge Headquarters	29.31, -89.86
East Grand Terre	San Bernard NWR (Big Pond Unit)	29.31, -89.86
East Grand Terre	Scenic Galveston	29.31, -89.86
East Grand Terre	Wisner	29.31, -89.86
Exxon Fields	Bolivar_Flats	29.25, -89.97
Exxon Fields	Brazoria National Wildlife Refuge	29.25, -89.97
Exxon Fields	East Grand Terre	29.25, -89.97
Exxon Fields	Exxon Fields	29.25, -89.97
Exxon Fields	Grand_Isle	29.25, -89.97
Exxon Fields	High_Island	29.25, -89.97
Exxon Fields	Quintana	29.25, -89.97
Exxon Fields	Rock W	29.25, -89.97
Exxon Fields	Rockport Bay Education Center	29.25, -89.97
Exxon Fields	Sabine NWR	29.25, -89.97

Appendix C

Appendix - Site Visits

C.1 Unique Visits

The following table shows the length in minutes of each unique visit:

C.2 Visits by tag and station

For each year and tag, the following table shows the total number and mean length (mins) of visits at each station:

C.3 Visits by Station

For each year and receiver, the following table shows the mean number and length (mins) of visits of all tags.

Table C.1: Time and length of each unique visit for each tag

Motus Tag ID	Lotek ID	Receiver	Location	Distance from Exxon F
23270	136	Quintana	28.93, -95.31	32
23216	91	Wisner	29.12, -90.16	1
23216	91	Wisner	29.12, -90.16	1
23216	91	Wisner	29.12, -90.16	1

Table C.2: Total number and mean length of visits (mins) for each tag by receiver

Year	Motus Tag ID	Lotek ID	Receiver	Total Number of Visits	Mean
2017	23209	84	Exxon Fields	6	
2017	23209	84	Grand_Isle	3	
2017	23212	87	Exxon Fields	36	
2017	23212	87	Grand_Isle	7	
2017	23214	89	Exxon Fields	113	
2017	23214	89	Grand_Isle	25	
2017	23216	91	Exxon Fields	94	
2017	23216	91	Grand_Isle	37	
2017	23216	91	Wisner	5	
2017	23217	92	Exxon Fields	30	
2017	23217	92	Grand_Isle	3	
2017	23218	93	East Grand Terre	2	
2017	23218	93	Exxon Fields	317	
2017	23218	93	Grand_Isle	120	
2017	23219	94	Exxon Fields	135	
2017	23219	94	Grand_Isle	18	
2017	23225	100	Exxon Fields	82	
2017	23225	100	Grand_Isle	15	
2017	23226	101	Exxon Fields	36	
2017	23226	101	Grand_Isle	3	
2017	23227	102	Exxon Fields	31	
2017	23227	102	Grand_Isle	2	
2017	23245	121	Exxon Fields	59	
2017	23245	121	Grand_Isle	13	
2017	23246	122	Exxon Fields	78	
2017	23246	122	Grand_Isle	24	
2017	23248	124	East Grand Terre	2	
2017	23248	124	Exxon Fields	33	
2017	23248	124	Grand_Isle	9	
2017	23249	126	Exxon Fields	13	
2017	23249	126	Grand_Isle	2	
2017	23251	129	Exxon Fields	109	
2017	23251	129	Grand_Isle	12	
2017	23251	129	Wisner	1	
2017	23257	135	East Grand Terre	1	
2017	23257	135	Exxon Fields	184	
2017	23257	135	Grand_Isle	45	
2017	23258	137	Exxon Fields	16	
2017	23258	137	Grand_Isle	1	
2017	23259	138	East Grand Terre	1	
2017	23259	138	Exxon Fields	22	
2017	23259	138	Grand_Isle	3	
2017	23261	141	Exxon Fields	158	
2017	23261	141	Grand_Isle	26	
2017	23261	141	Wisner	12	
2017	23262	142	Exxon Fields	83	
2017	23262	142	Grand_Isle	7	
2017	23270	136	Exxon Fields	67	
2017	23270	136	Grand_Isle	7	
2017	23270	136	Quintana	1	
2017	23270	136	San Bernard National Wildlife Refuge Headquarters	1	
2017	23270	136	San Bernard NWR (Big Pond Unit)	2	

Table C.3: Mean length (mins) and number of visits of all tags at each receiver by year

Year	Receiver	Mean Length of Visits (mins)	Mean Number of Visits
2017	Exxon Fields	4.54	81.05
2017	Grand_Isle	2.52	18.19
2017	East Grand Terre	2.57	1.50
2017	Wisner	4.75	6.00
2017	Scenic Galveston	1.00	3.00
2017	Quintana	0.56	1.00
2017	San Bernard National Wildlife Refuge Headquarters	0.75	1.00
2017	San Bernard NWR (Big Pond Unit)	1.22	2.00
2018	Exxon Fields	4.63	78.79
2018	Grand_Isle	2.51	14.43
2018	East Grand Terre	4.20	3.18
2018	Bolivar_Flats	0.56	1.00