Exploratory plots

1 Summary of Data

1. Study area: Originally we planned for South California Air Basin (SCAB), but we refined to the area west to 117.75 W. We decide to drop data in Riverside because of different intended Wednesdays, as we show later.

2. Data source: From dr0328, we have four data source: MESA Agency (daily), MESA (2-week), Spiromics Agency (daily) and Spiromics (daily). We found that MESA Agency and Spiromics Agency are identical in SCAB. 2-week MESA and Spiromics are almost identical, except for multiple measurements from the same location with different site_type (F?L?O?). We decide to use only MESA data.

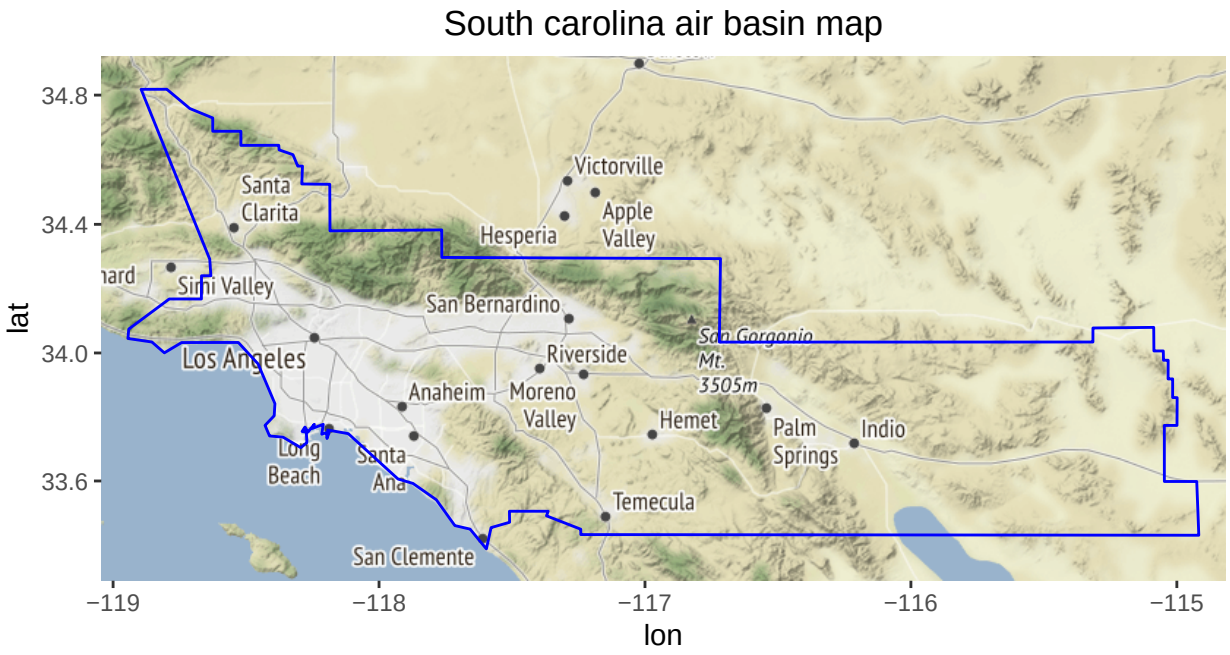
3. Time window: While the data are available from 1988-2018. We restrict our ST model to use data from 2000-2010.

4. Colocation: We found that one short term MESA site (mesa_L001) is likely to be colocated with a MESA Agency site. Thus, we will drop it in the analysis.

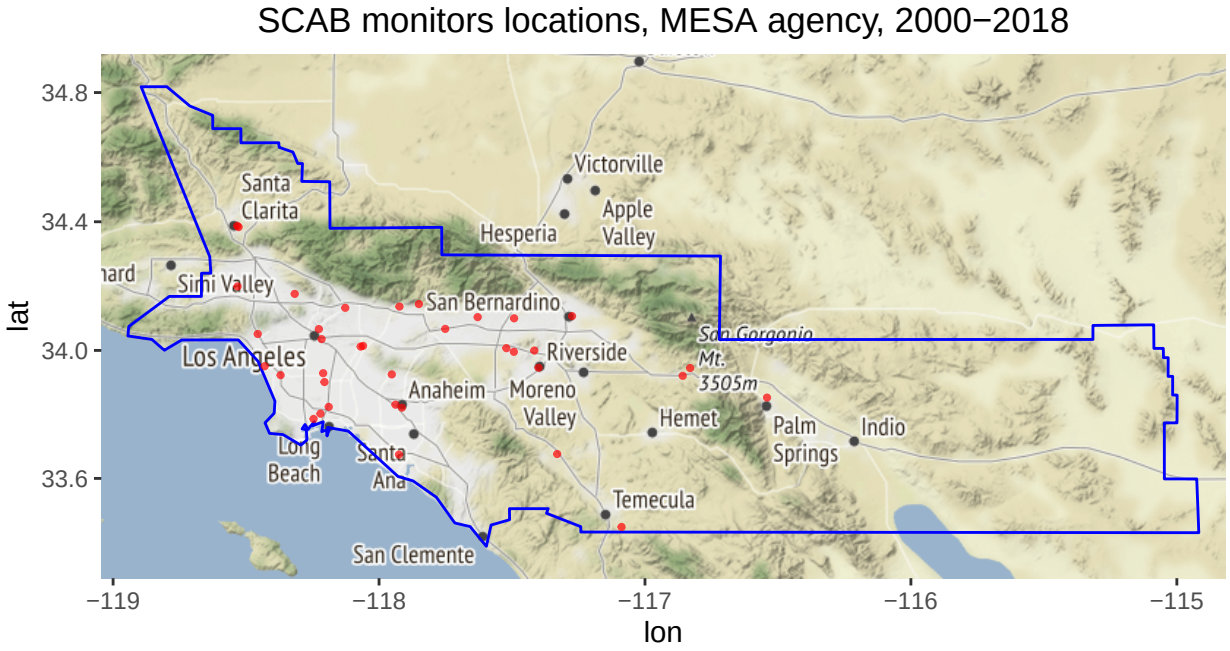
5. Intended Wednesday: For 2-week measures, we take the following approach to get the intended Wednesday (if not provided in the data set). We first find the mid-day of the start date and end date; then we assign the closest Wednesday to the mid-day as the intended Wednesday. In the end, we check which the two 2-week grid that the measurement is on. We do observe cases that measurements from the same site change from one grid to the other and we have not decided how to deal with it.

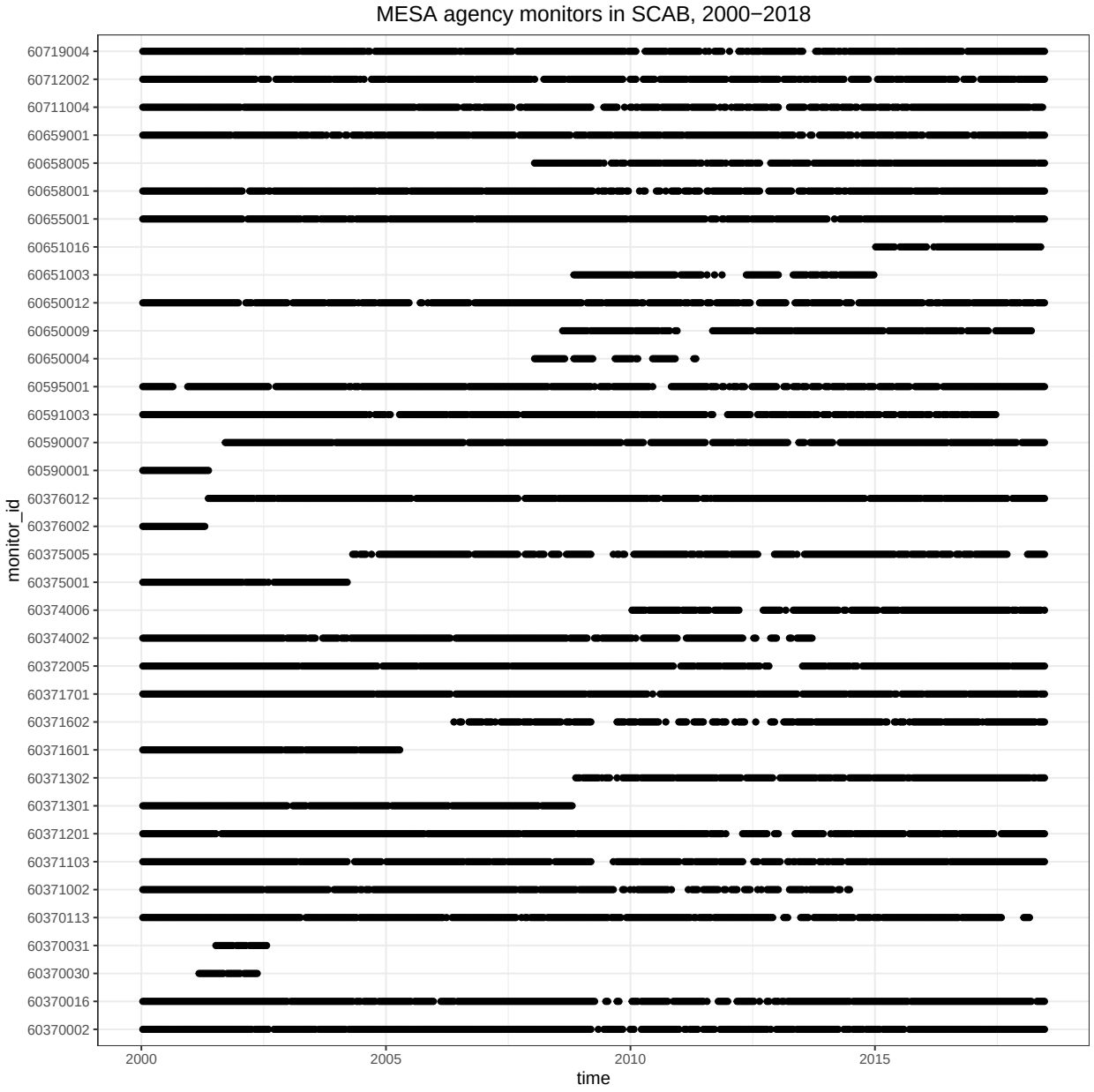
6. Daily measure to 2-week measure: For MESA Agency data, we want to aggregate them to form 2-week measures such that all data are on the same time resolution. We first choose a 2-week grid identified by a series of intended (middle) Wednesday. For each Wednesday, we check how many days that the site has measurements in this 2-week window centered by the intended Wednesday. If there are more than 12 measurements (>80% nonmissing), we then take the average of them and assign it as the 2-week measure. If more than 2 days are missing, we will treat this time window as missing.

2 South California Air Basin



3 MESA Agency



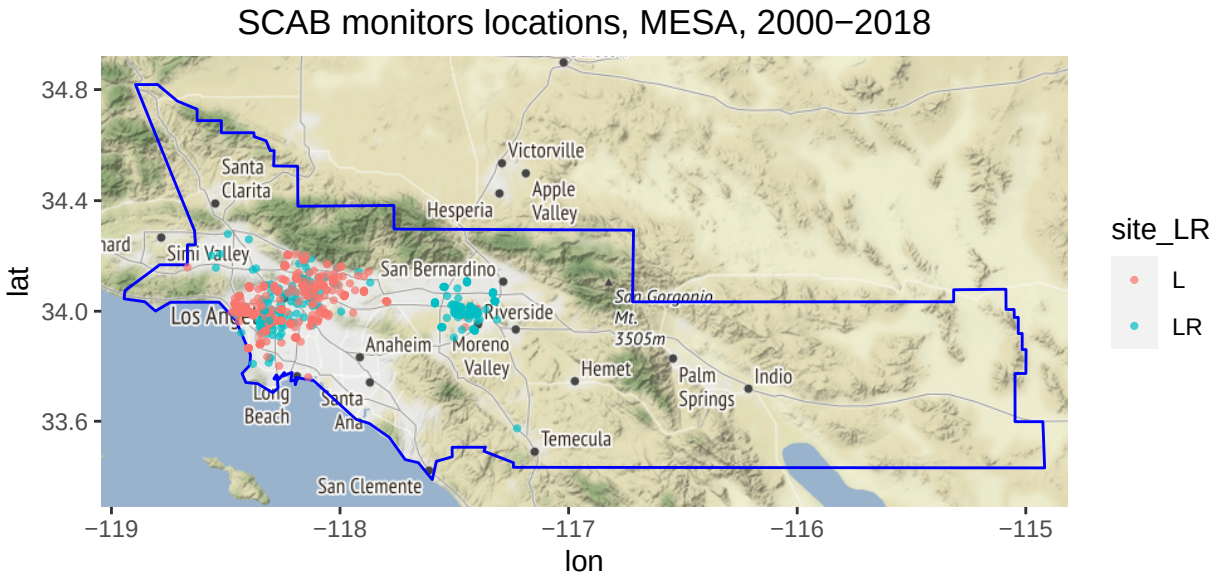

MESA agency monitors in SCAB, 2000−2018
monitor_id
60719004
60712002
60711004
60659001
60658005
60658001
60655001
60651016
60651003
60650012
60650009
60650004
60595001
60591003
60590007
60590001
60376012
60376002
60375005
60375001
60374006
60374002
60372005
60371701
60371602
60371601
60371302
60371301
60371201
60371103
60371002
60370113
60370031
60370030
60370016
60370002
2000
2005
2010
2015
time

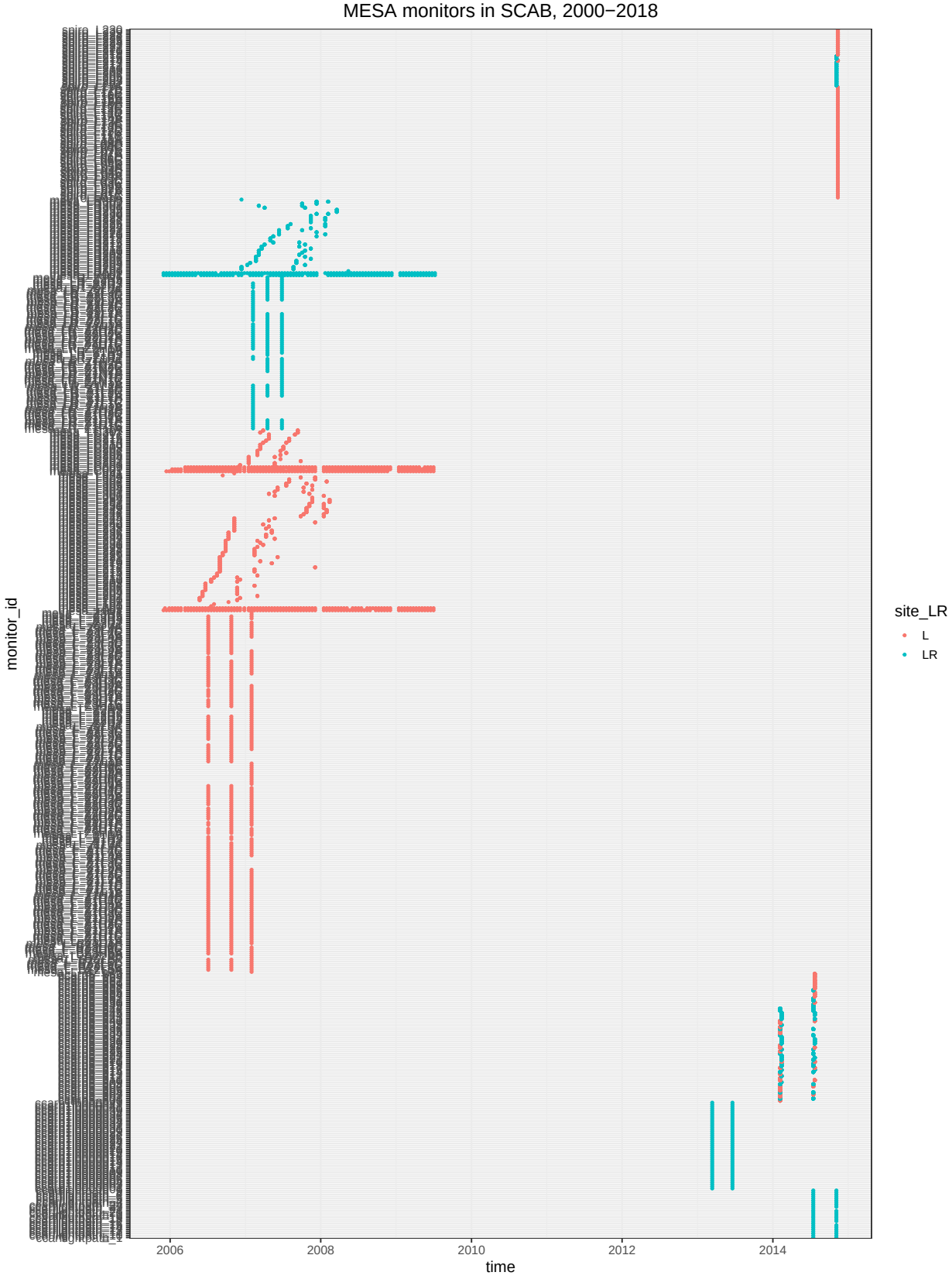
4 MESA

Some repitation. 2-4 rows all same except for 'site_type'.

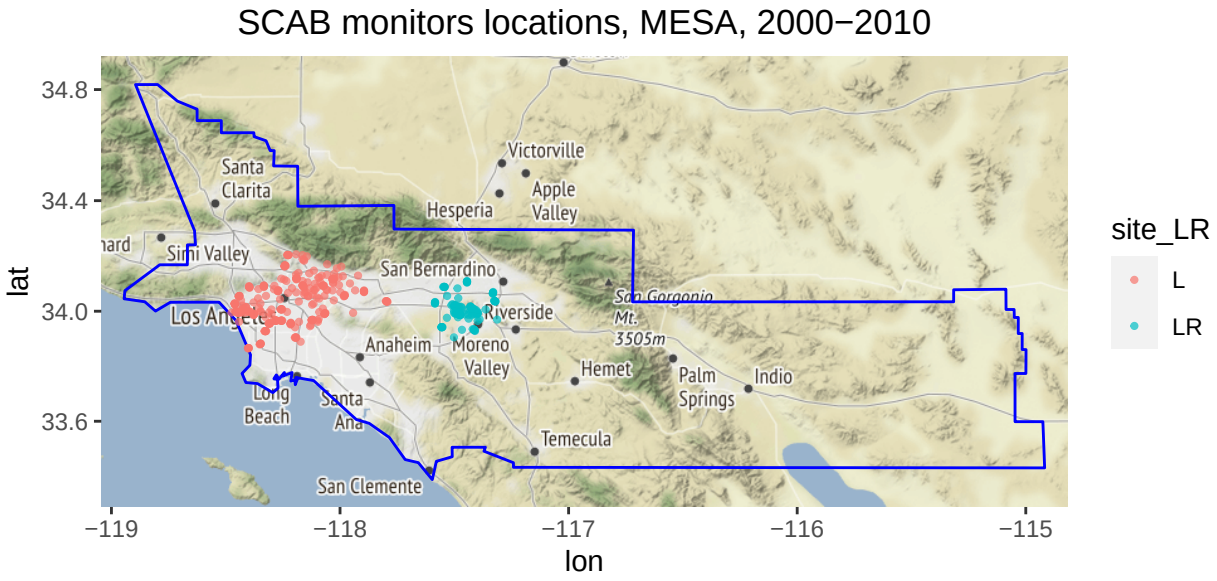
```
# add intended wednesday for missing
mesa_no2_scab$sample_start_date <- as.Date(mesa_no2_scab$sample_start_date)
mesa_no2_scab$sample_stop_date <- as.Date(mesa_no2_scab$sample_stop_date)

mesa_no2_scab$mid_day <-  mesa_no2_scab$sample_start_date +
   floor((mesa_no2_scab$sample_stop_date-mesa_no2_scab$sample_start_date)/2)

mesa_no2_scab$candidate_wed <- round_date(mesa_no2_scab$mid_day, "week",3)

mesa_no2_scab <- as.data.table(mesa_no2_scab)

mesa_no2_scab <- mesa_no2_scab[is.na(intended_wednesday)|intended_wednesday == "",
                               ind := 'Missing']

mesa_no2_scab$intended_wednesday <- as.Date(mesa_no2_scab$intended_wednesday)

mesa_no2_scab <- mesa_no2_scab[ind == 'Missing',
                               intended_wednesday := candidate_wed]


# classify by intended wednesdays

mesa_no2_scab$wed_week <- (as.numeric(mesa_no2_scab$intended_wednesday -as.Date("1998-12-09"))%% 14)/7

mesa_no2_scab <- mesa_no2_scab[wed_week == 1,
                               wed_type := 'L']
mesa_no2_scab <- mesa_no2_scab[wed_week == 0,
                               wed_type := 'LR']

mesa_no2_scab <- mesa_no2_scab [ ,site_LR := ifelse(mean(wed_week)>0.5,'L','LR'), by=native_id]
```

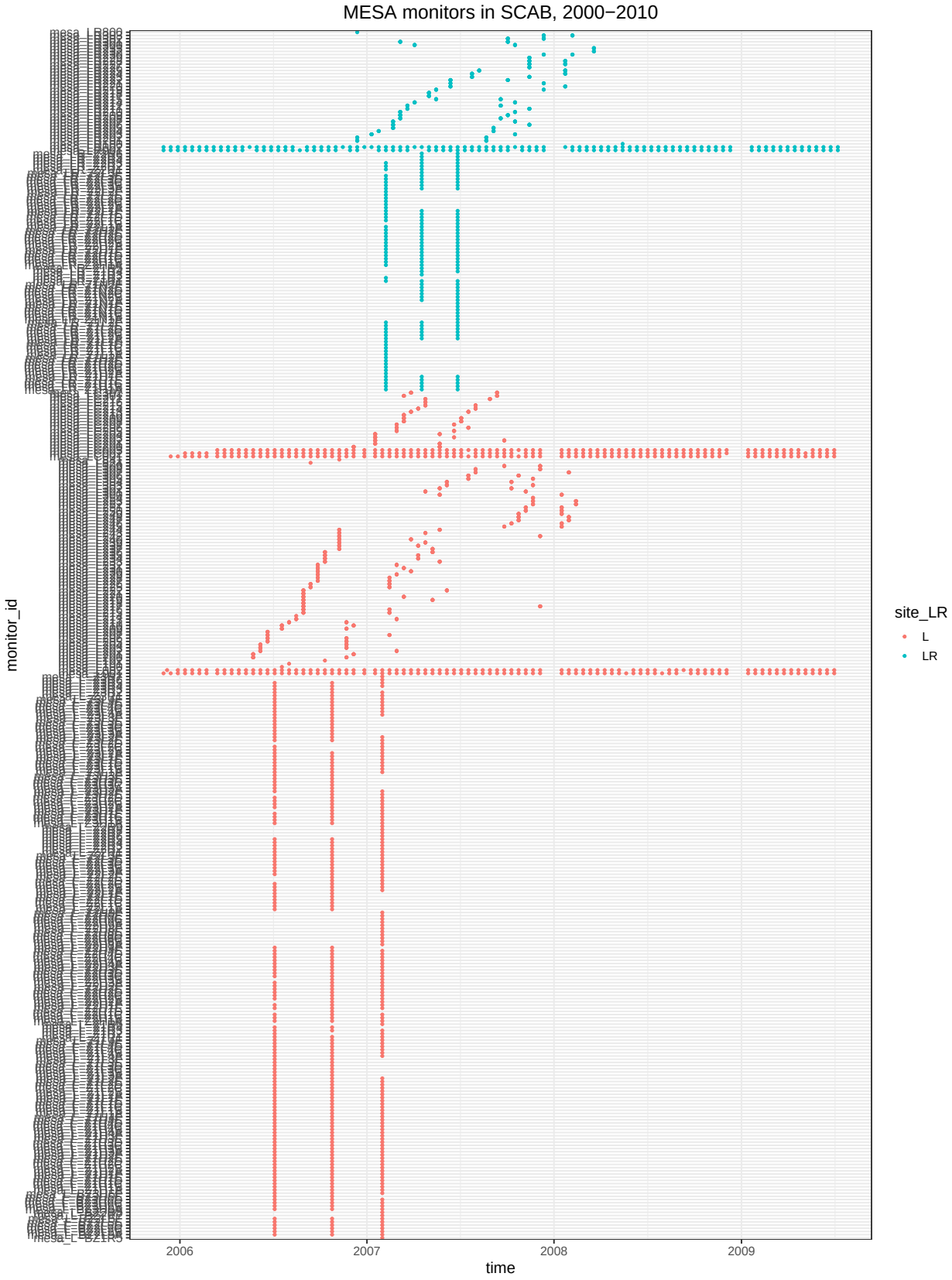
4.1 MESA long term by L and LR



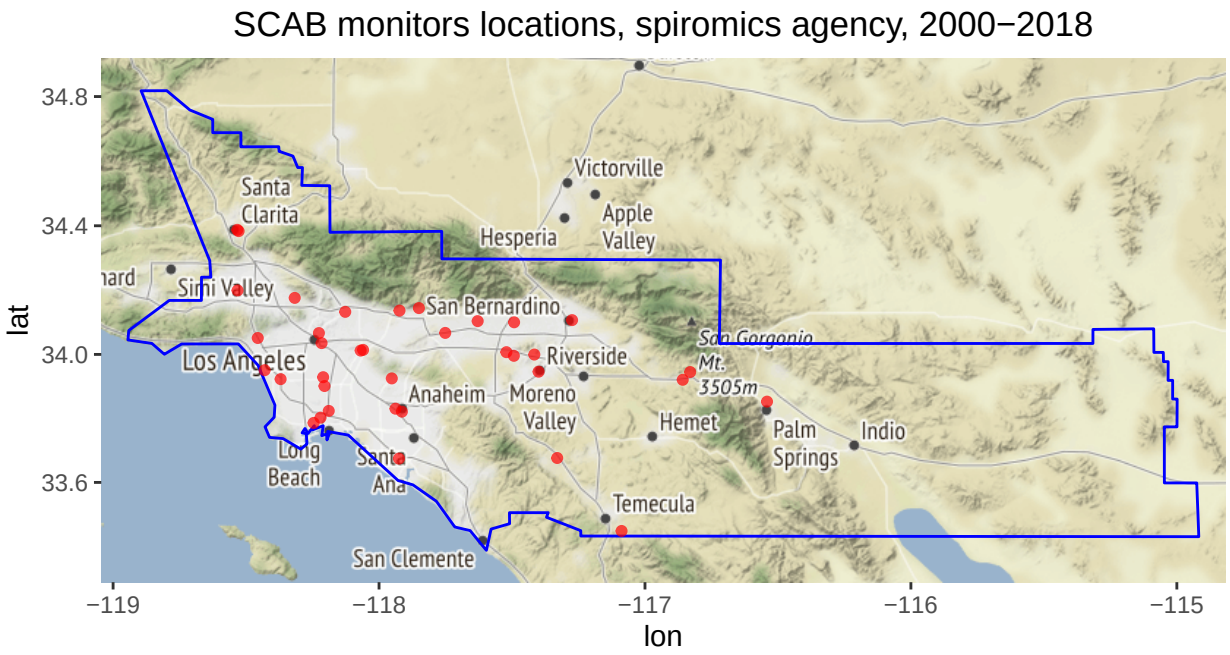

MESA monitors in SCAB, 2000−2018
monitor_id
2006
2008
2010
2012
2014
time
site_LR
L
LR

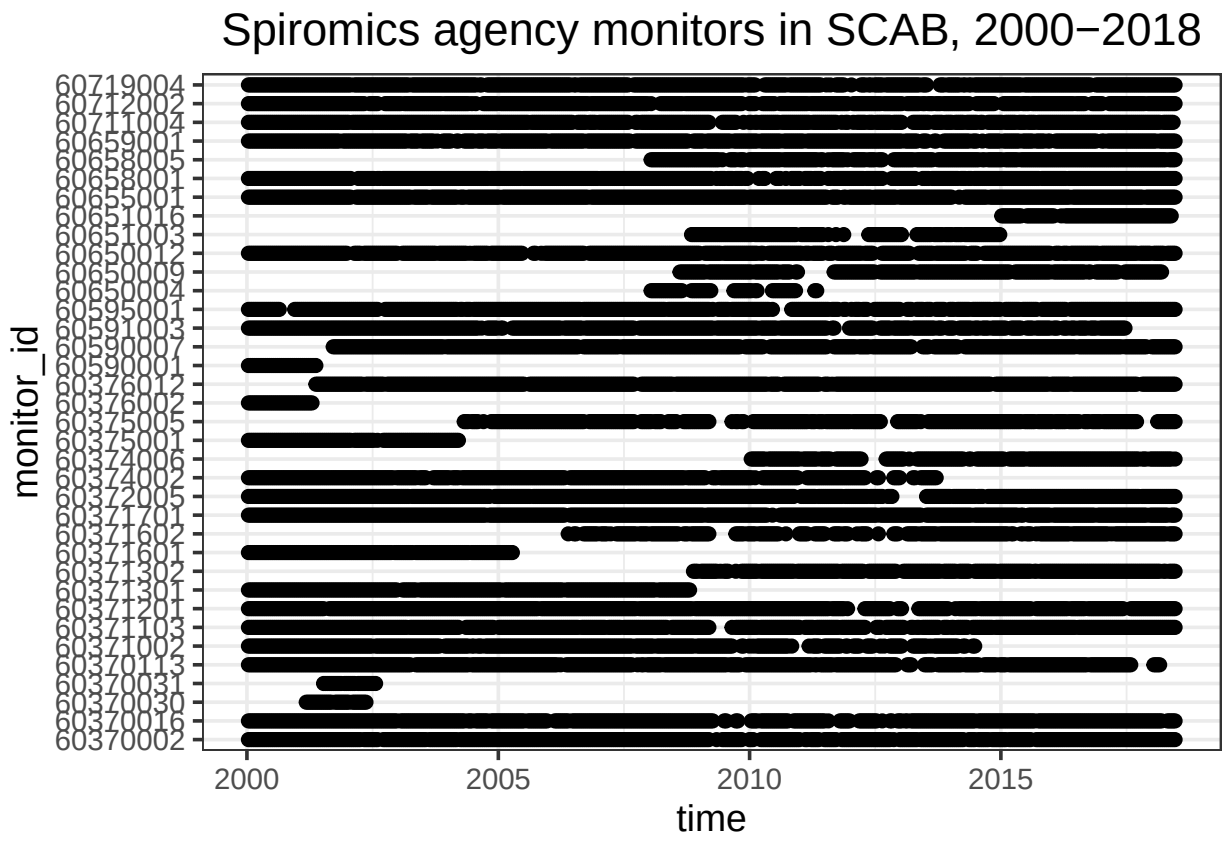
4.2 MESA short term by L and LR




MESA monitors in SCAB, 2000−2010
monitor_id
time
2006
2007
2008
2009
site_LR
L
LR

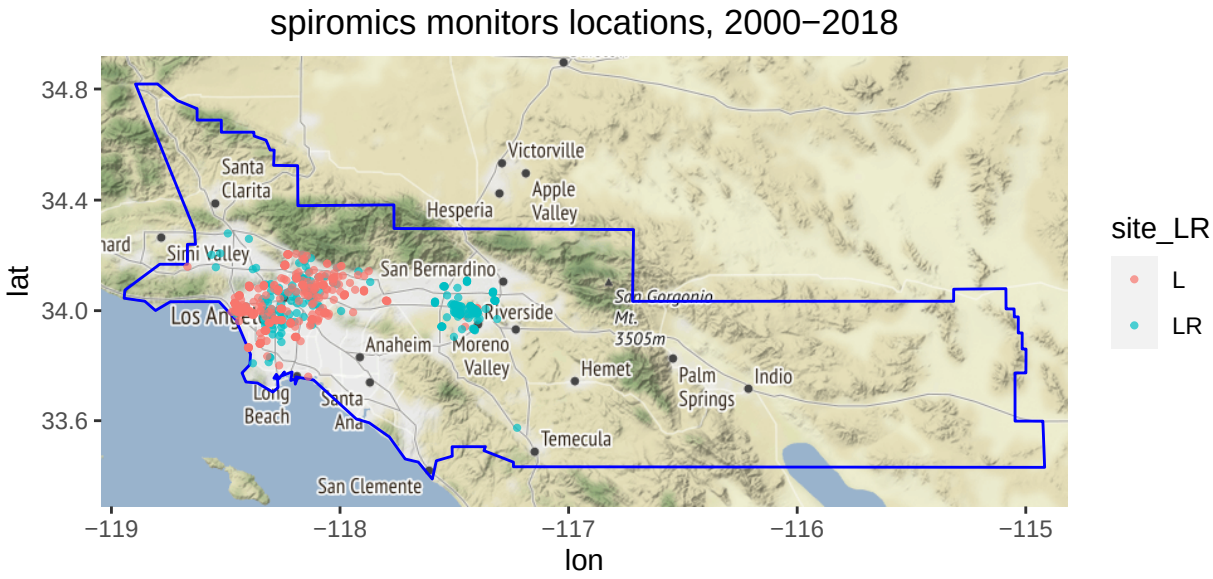
5 Spiromics Agency

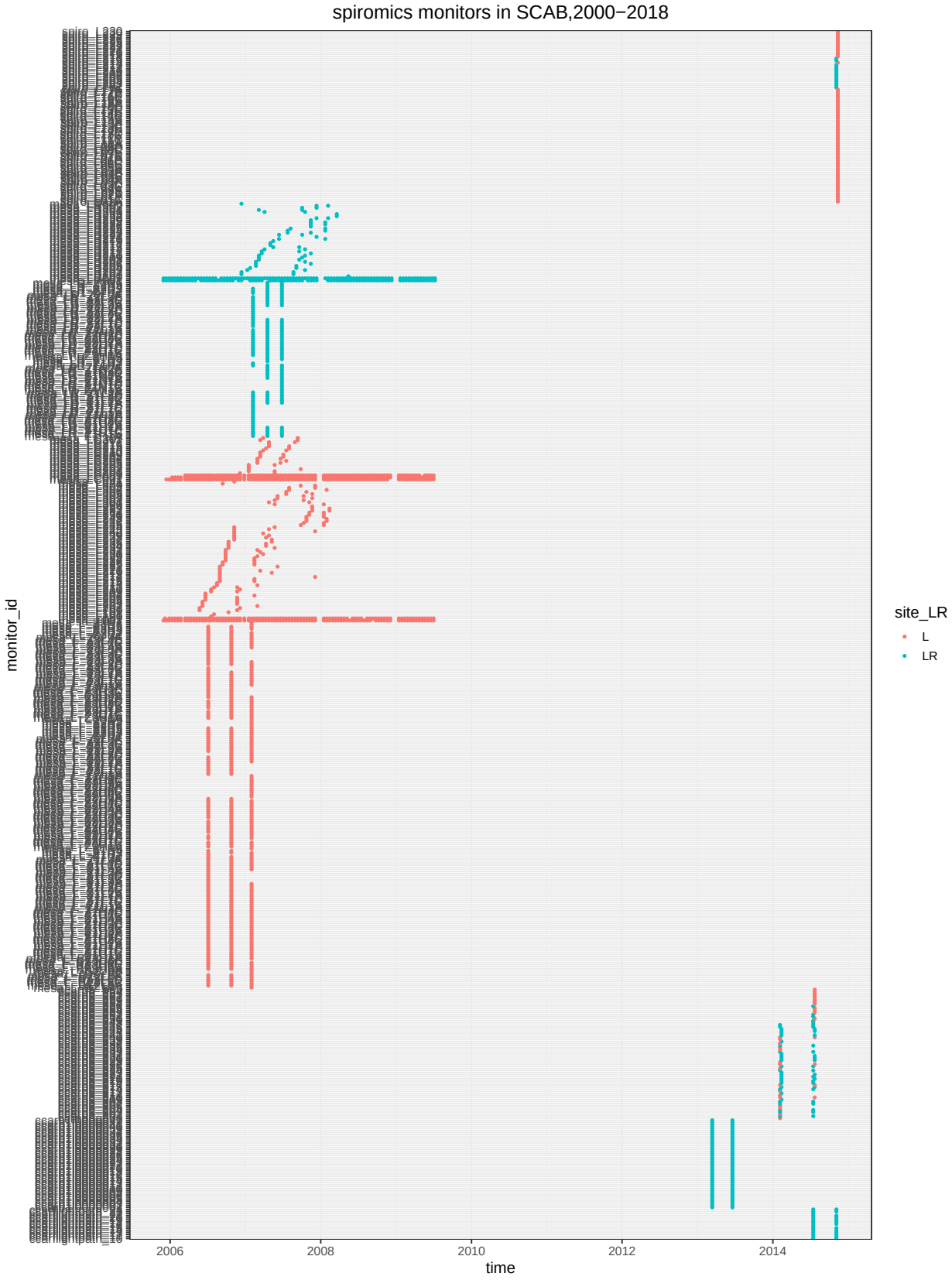



Spiromics agency monitors in SCAB, 2000−2018
monitor_id
60719004
60712002
60711004
60659001
60658005
60658001
60655001
60651016
60651003
60650012
60650009
60650004
60595001
60591003
60590007
60590001
60376012
60376002
60375005
60375001
60374006
60374002
60372005
60371701
60371602
60371601
60371302
60371301
60371201
60371103
60371002
60370113
60370031
60370030
60370016
60370002
2000
2005
2010
2015
time

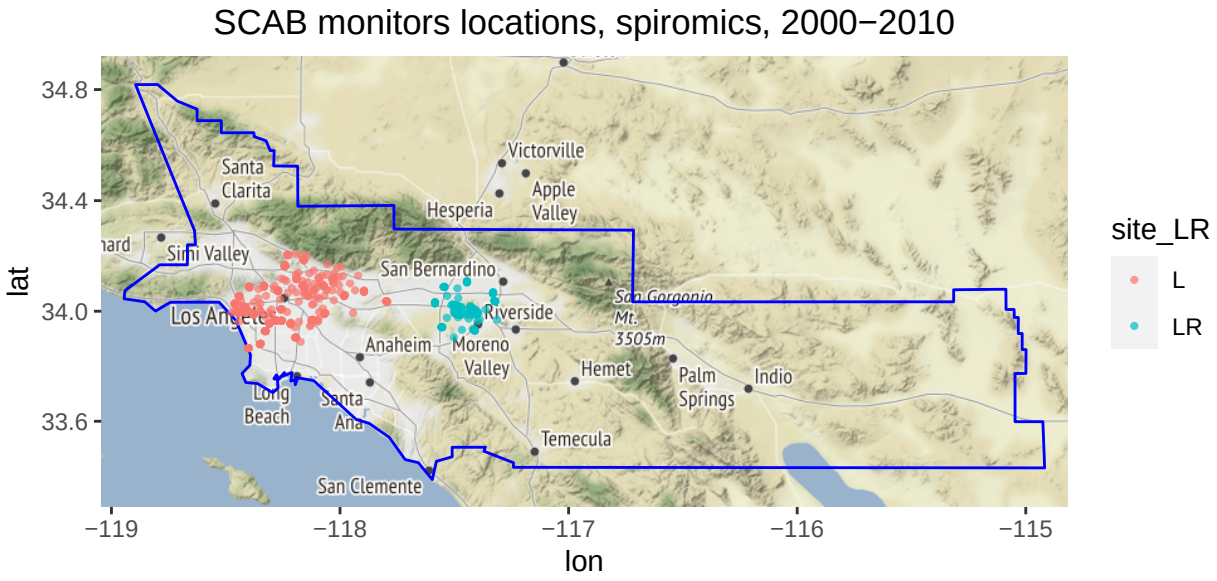
6 Spiromics

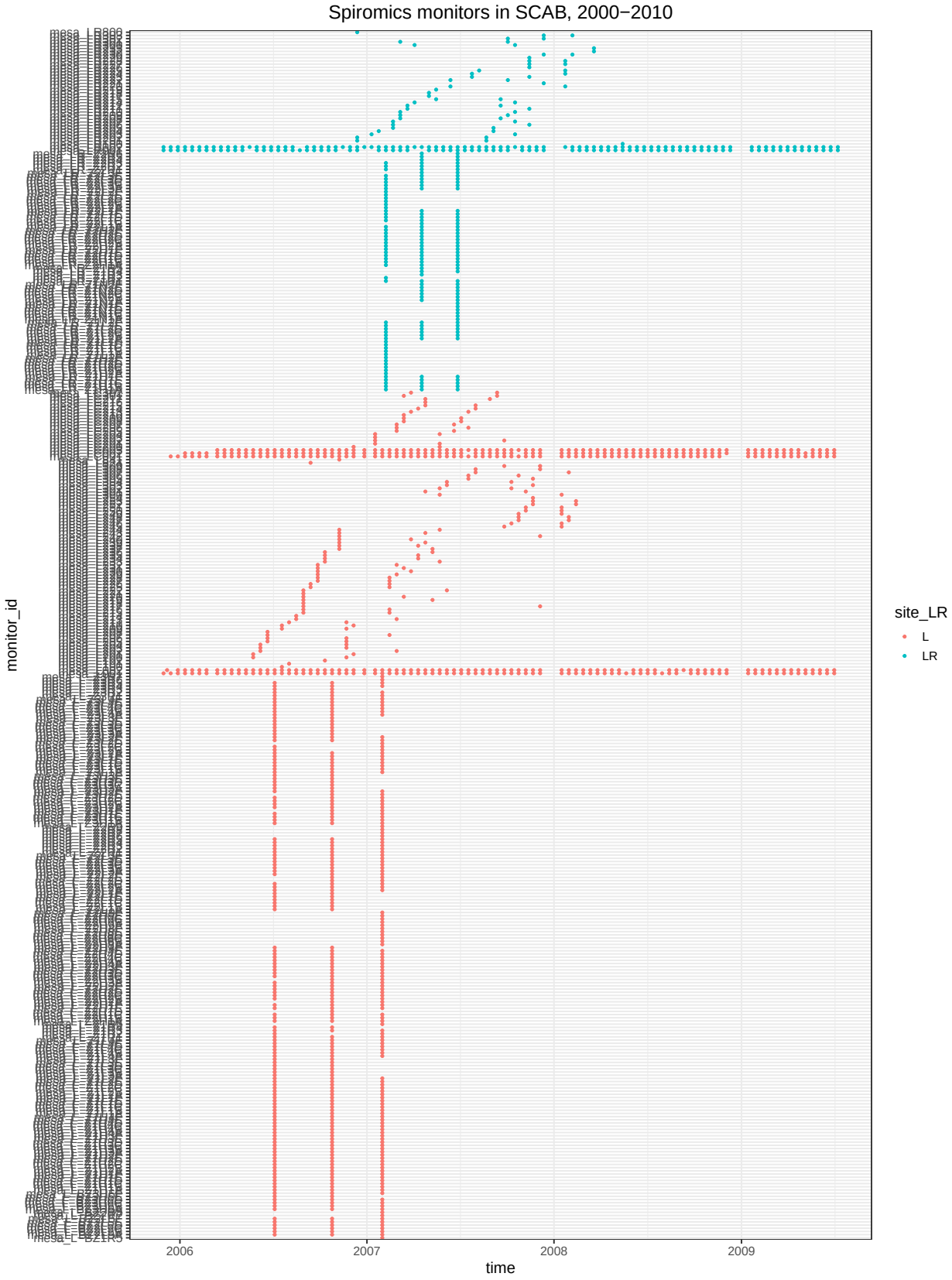
6.1 spiromics long term by L and LR




spiromics monitors in SCAB,2000−2018
monitor_id
time
2006
2008
2010
2012
2014
site_LR
L
LR

6.2 spiromics early term by L and LR




Spiromics monitors in SCAB, 2000−2010
monitor_id
2006
2007
2008
2009
time
site_LR
L
LR