Data Analysis Sensitivities

Minimum Detection Limits and Data Completeness

Elizabeth Mannshardt, US EPA

Outline

Data Reporting and Record Handling Considerations

- MDL Sensitivity

- Completeness Sensitivity

- Discussion

MDL

Sensitivity Analysis for MDL Decisions

MDL Completeness

MDL considerations of which sites to include in trends:

- Exclude site-pollutant-years if a certain percent data is below MDL when being used in aggregating across sites. If more than 65% of data reported were below the MDL for at least half of the years of a given trend, the trend was removed.

- For example - a site with four annual averages with 75% of data below the MDL would be removed. In contrast, a site with two annual averages with 100% of data below MDL would be included, as long as less than 65% of data were below the MDL for the other years.

UrbanToxicTrendsDocumentation_TechMemo_Dec2011.pdf
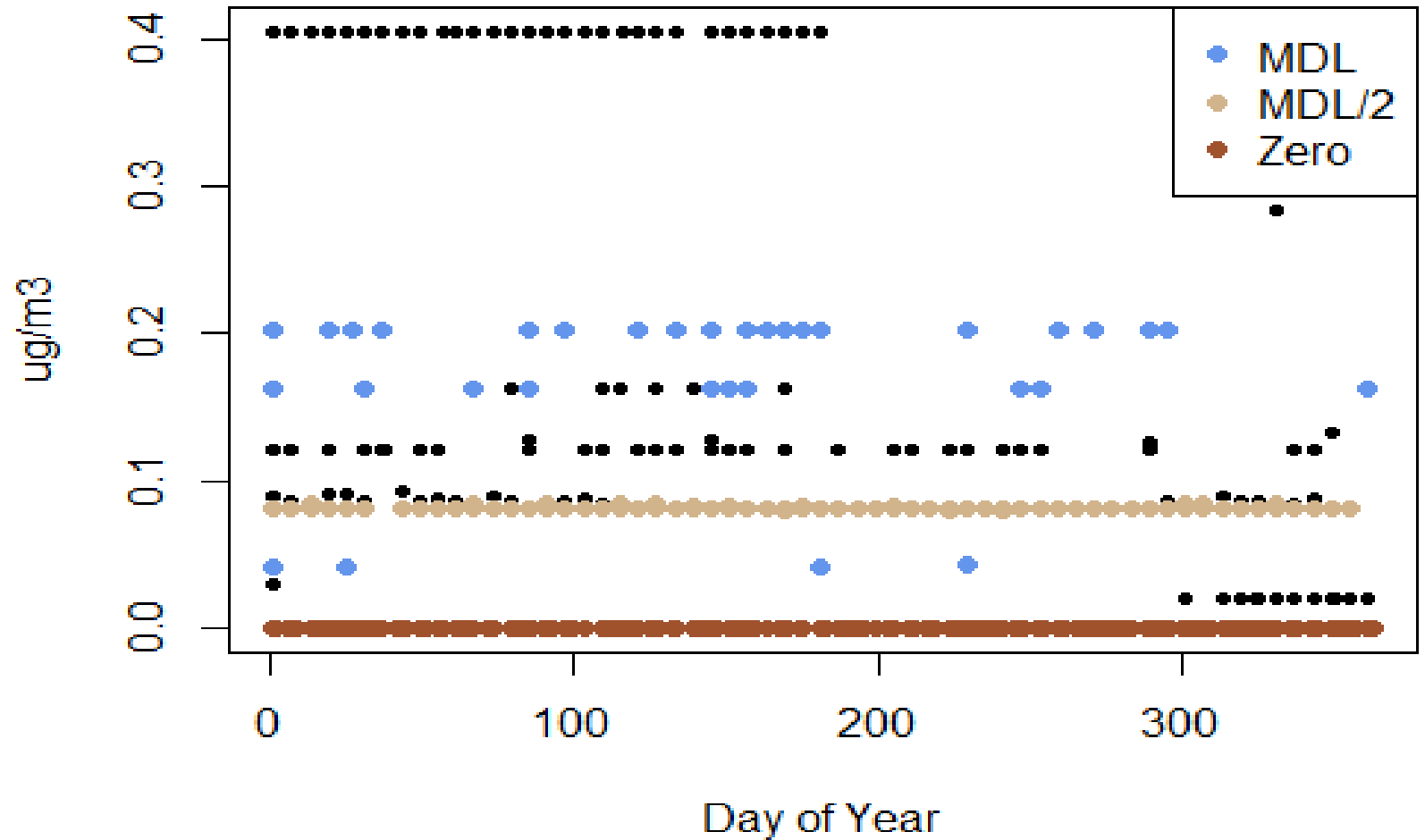
MDL Sensitivity Analysis

Objective: Assess how different non-detect substitution algorithms would affect overall trends.

Investigate four specific algorithms:

- utilizing observations below MDL as recorded
- replacing non-detects with zero
- replacing non-detects with one-half the detection limits (MDL/2) — an algorithm widely used in environmental health evaluations
- replacing all non-detects with NA (removing from analysis)

Ethylene Dichloride: Below MDL

Ethylene dichloride 1992 <= MDL

In 1992, 2407 out of 2701 observations below the MDL (89%)
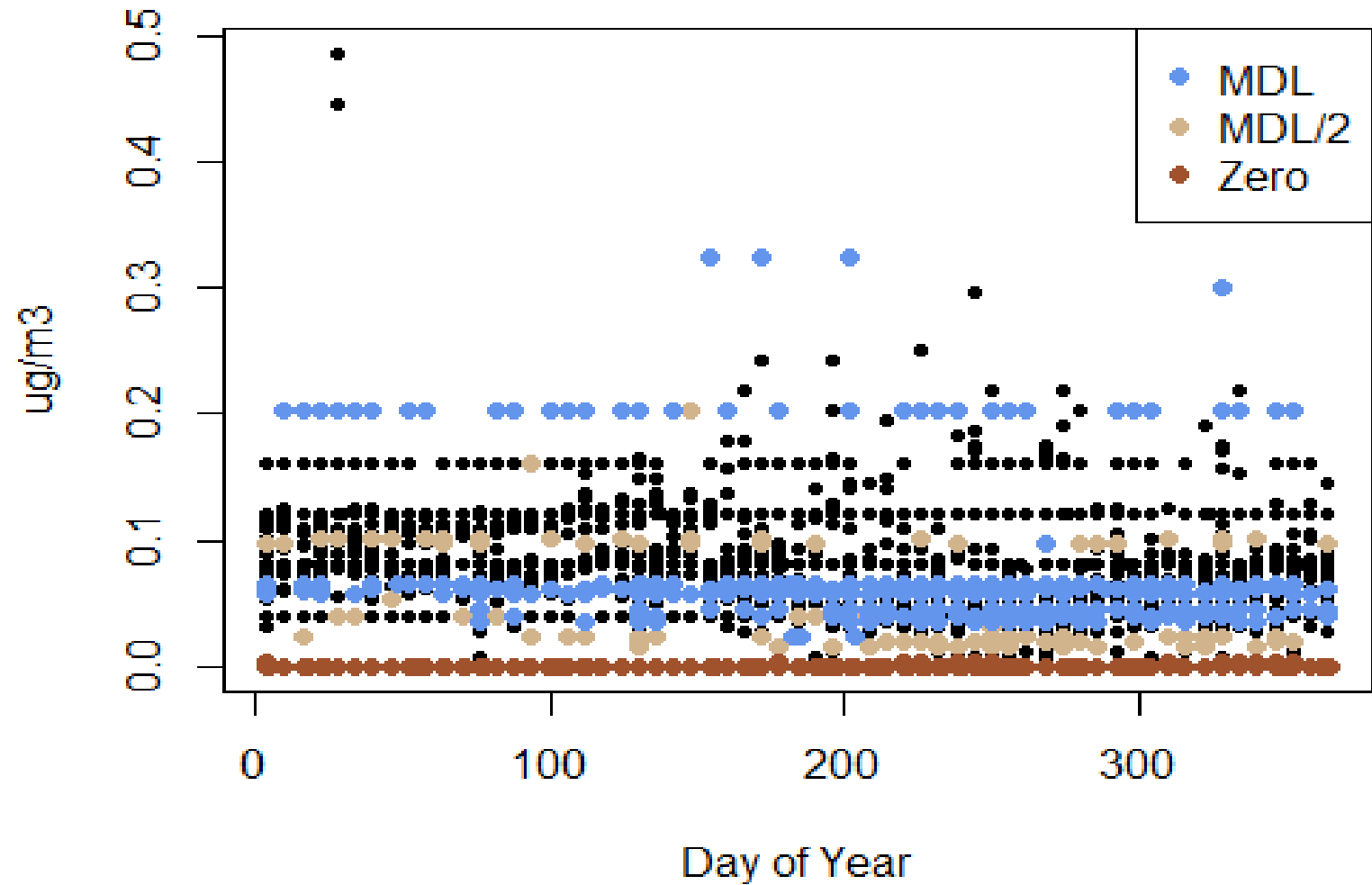
MDL Sensitivity

Annual averages are considered utilizing the replacement schemes for observations below the MDL

Ethylene Dichloride Averages	ALL	<MDL-> MDL	<MDL-> MDL/2	<MDL-> Zero	<MDL-> NA
1992	0.19	0.41	0.29	0.16	1.48

The range produced by different MDL methods is up to **1.32** with a percent difference as much as **819%.**

Ethylene Dichloride: Below MDL

Ethylene dichloride 2013 <= MDL

Legend: MDL, MDL/2, Zero

y-axis: ug/m3

x-axis: Day of Year

In 2013, 7464 out of 11622 observations were below the MDL (64%)

MDL Sensitivity

Ethylene Dichloride Averages	ALL	<MDL-> MDL	<MDL-> MDL/2	<MDL-> Zero	<MDL-> NA
2013:	0.095	0.210	0.143	0.075	0.210

The range produced by different MDL methods is up to **0.14** with a percent difference as much as **181%**.

Benzene: Below MDL

Benzene 2013 <= MDL
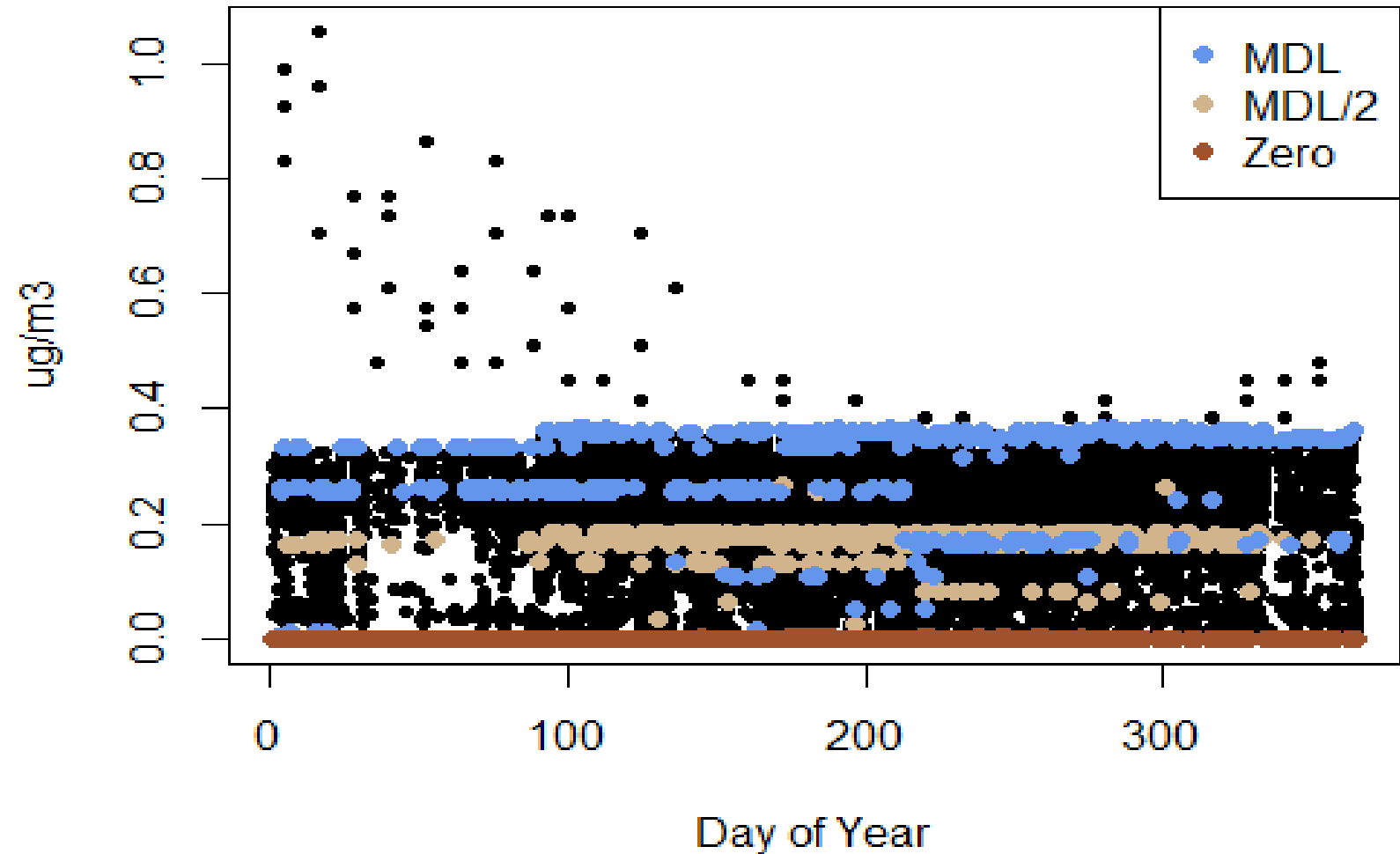
In 2013, 32,692 out of 182,355 observations were below the MDL (18%)

MDL Sensitivity

Benzene	ALL	>MDL->MDL	>MDL->MDL/2	>MDL->Zero	>MDL->NA
1992	3.14	3.18	3.16	3.13	3.44
2013:	0.71	0.73	0.71	0.69	0.85

For 1992, the range produced by different MDL methods is **0.307** with a percent difference as much as **10%**.

For 2013, the range produced by different MDL methods is **0.152** with a percent difference as much as **22%**.

Take Away Message

- Replacing MDL's has a significant effect on trends analysis

- Caution must be used

Completeness

Sensitivity Analysis for Data Completeness Decisions

Completeness

Completeness considerations as to which sites to include in trends:

Considerations for this analysis:
- Quantity of data – how many measurements are needed within a year to get annual statistics
- Seasonal representativeness
- Seasonal bias

Other considerations:
- For trends—do sites have sufficient years of data

Ie, for a 5 year trend, do you need at least 12 years (75% completeness) for any given site?
- MDL handling (% above MDL)
- Etc

Observations by Site

1,3-Butadiene 2013

Benzene 2013

Ethylene dichloride 2013

Formaldehyde 2013
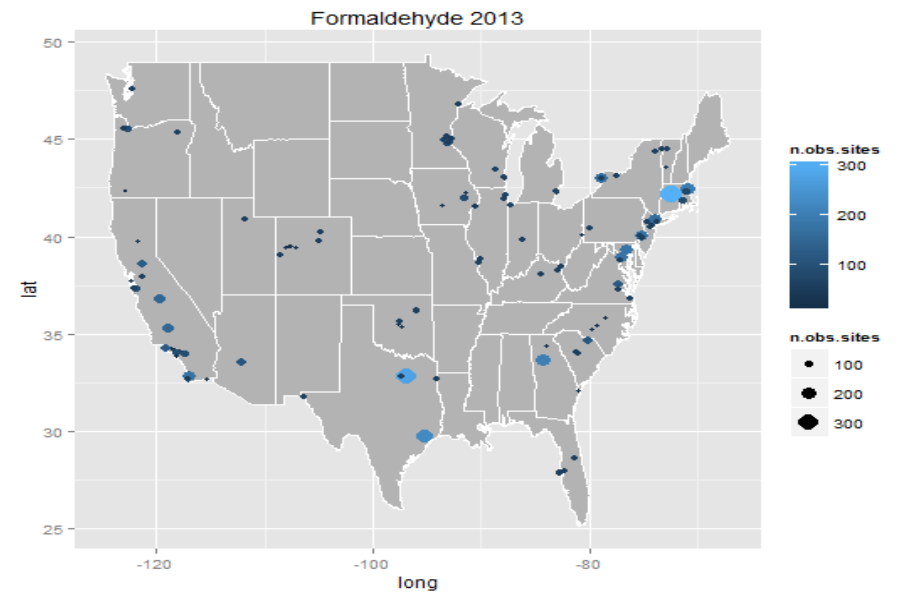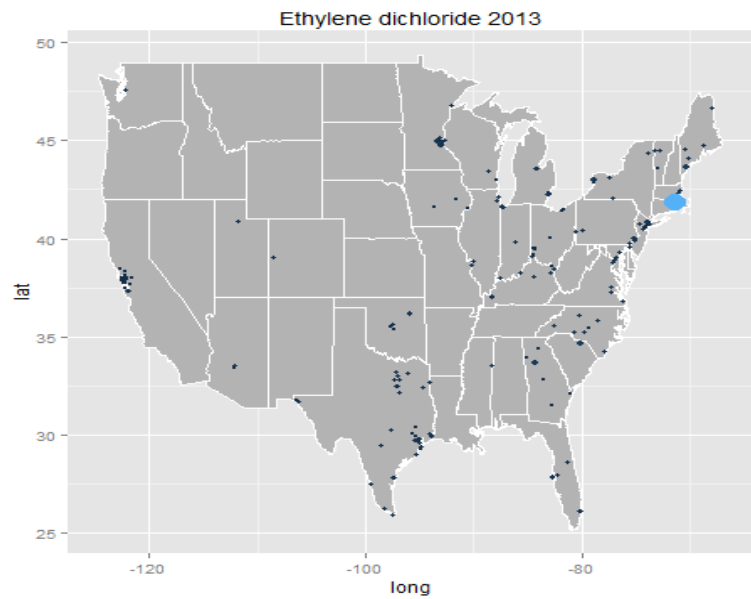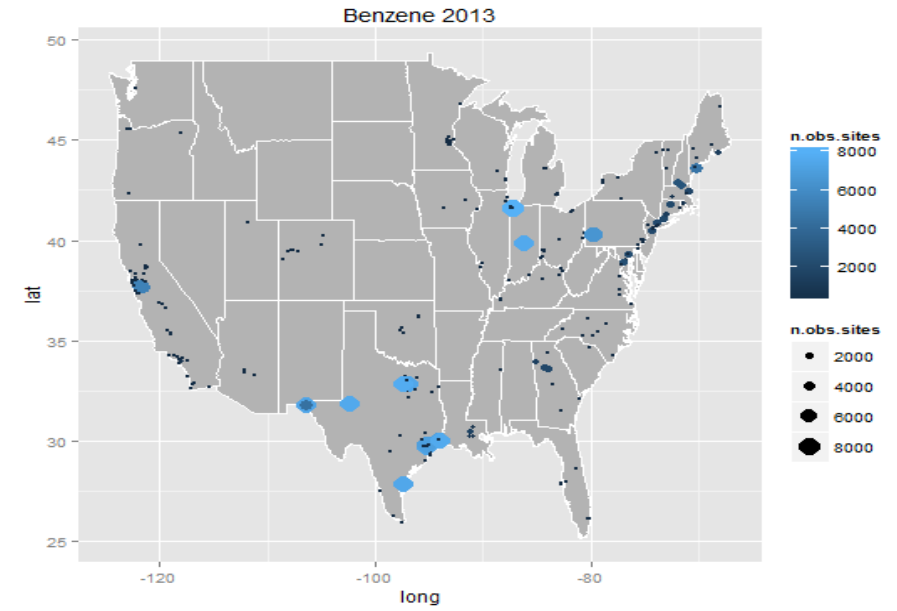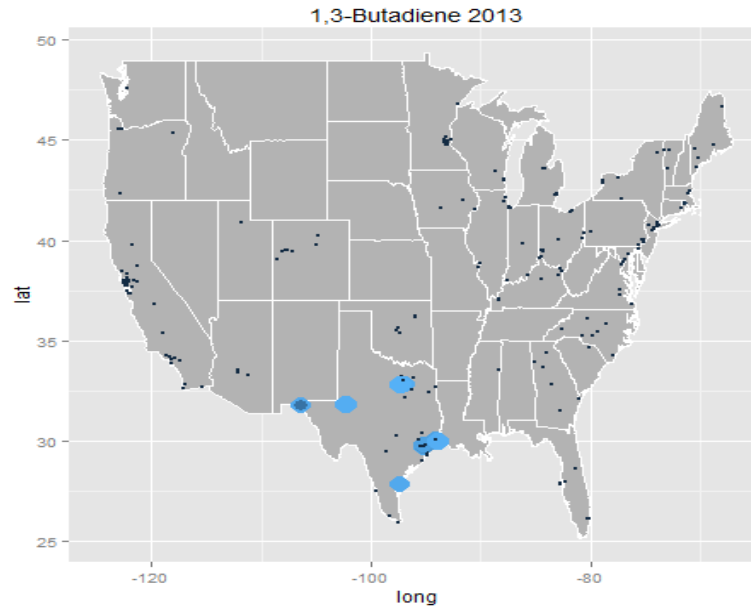
Completeness Sensitivity Analysis

Objective: Assess how different completeness requirements affect overall trends.

Investigate completeness scenarios

- Consider completeness for 6-day versus 12-day measurements for annual trend
- Consider seasonal representativeness

Data per Site by Month

	1,3-Butadiene	Benzene	Ethylene Dichloride	Formaldehyde
	Houston, TX	Houston, TX	Rumford, RI	Dallas, TX
Jan	618	618	6	5
Feb	517	517	10	5
Mar	646	646	10	5
Apr	624	624	8	5
May	567	567	4	5
Jun	637	637	4	77
Jul	673	673	5	80
Aug	664	664	7	80
Sept	520	519	8	4
Oct	645	644	10	6
Nov	656	656	10	5
Dec	678	677	10	4
Total	7442	7445	92	281

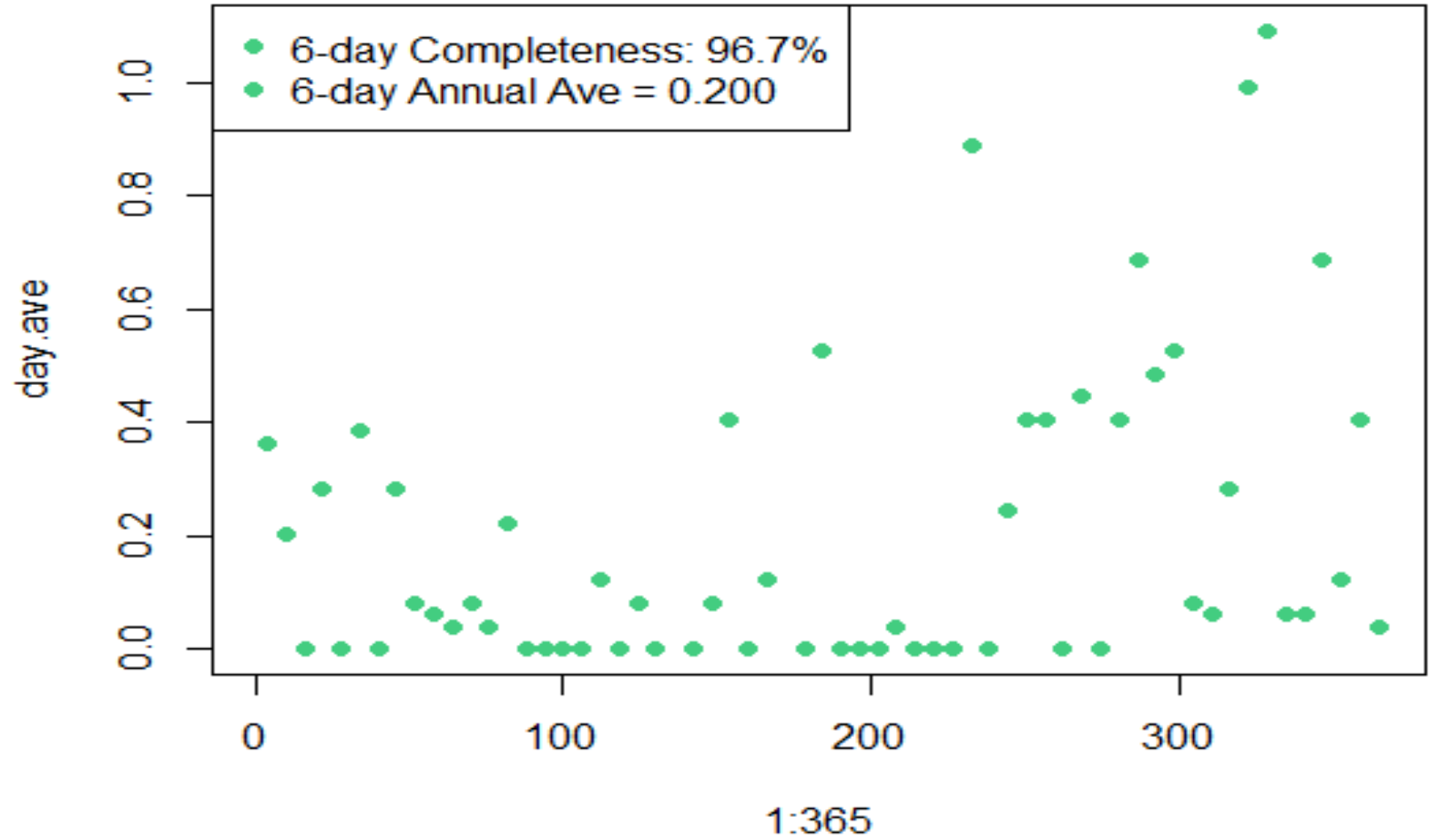Table of observations per toxic by month

Ethylene Dichloride

Complete over time on a 6-day scale

- Consider average across 6-day measurements

- Consider a >75% complete (but not seasonally complete) analysis versus a 12-day composite analysis that is 49% complete and evenly distributed over time
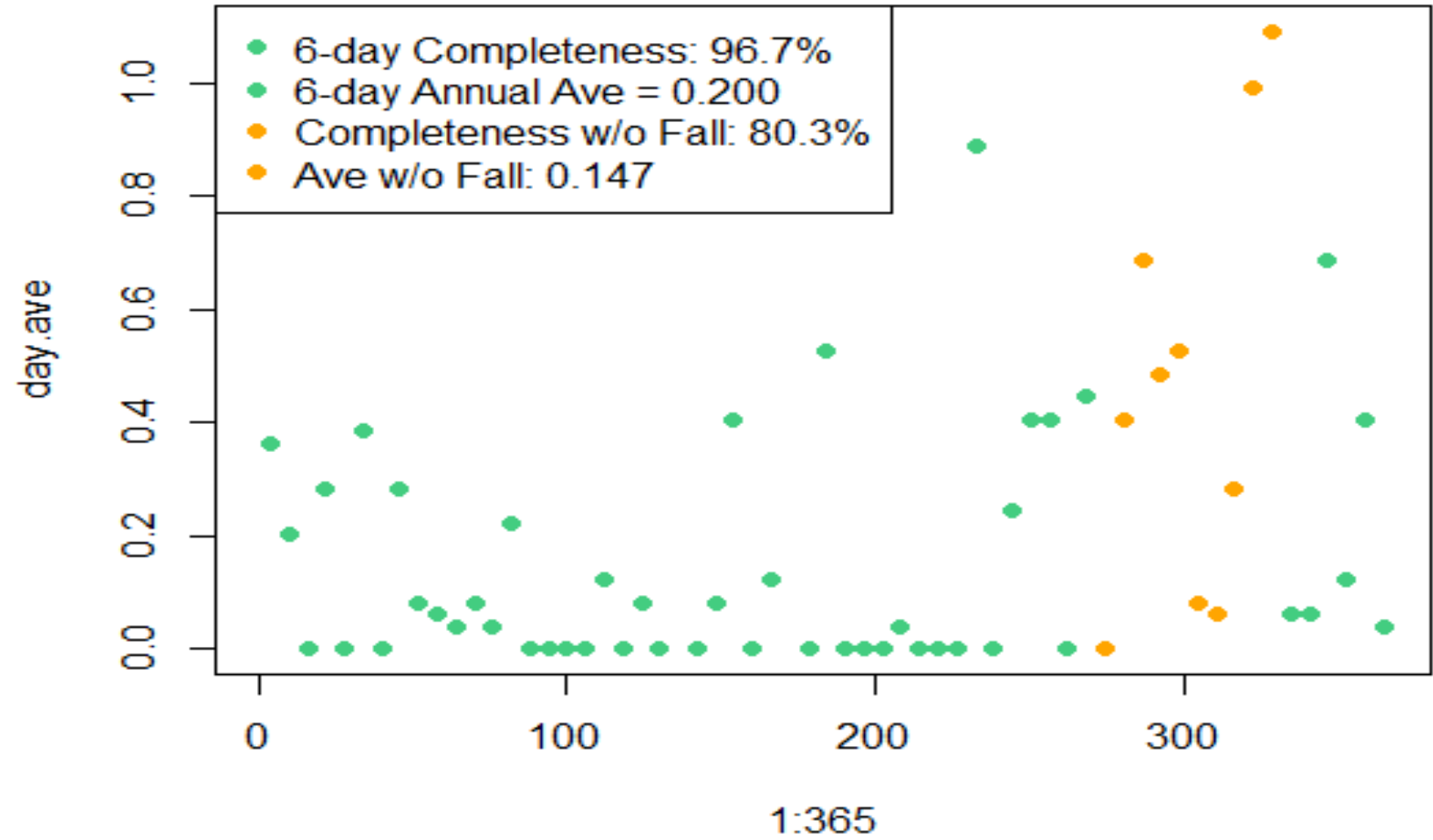
- Compare effect on the averages

Ethylene Dichloride
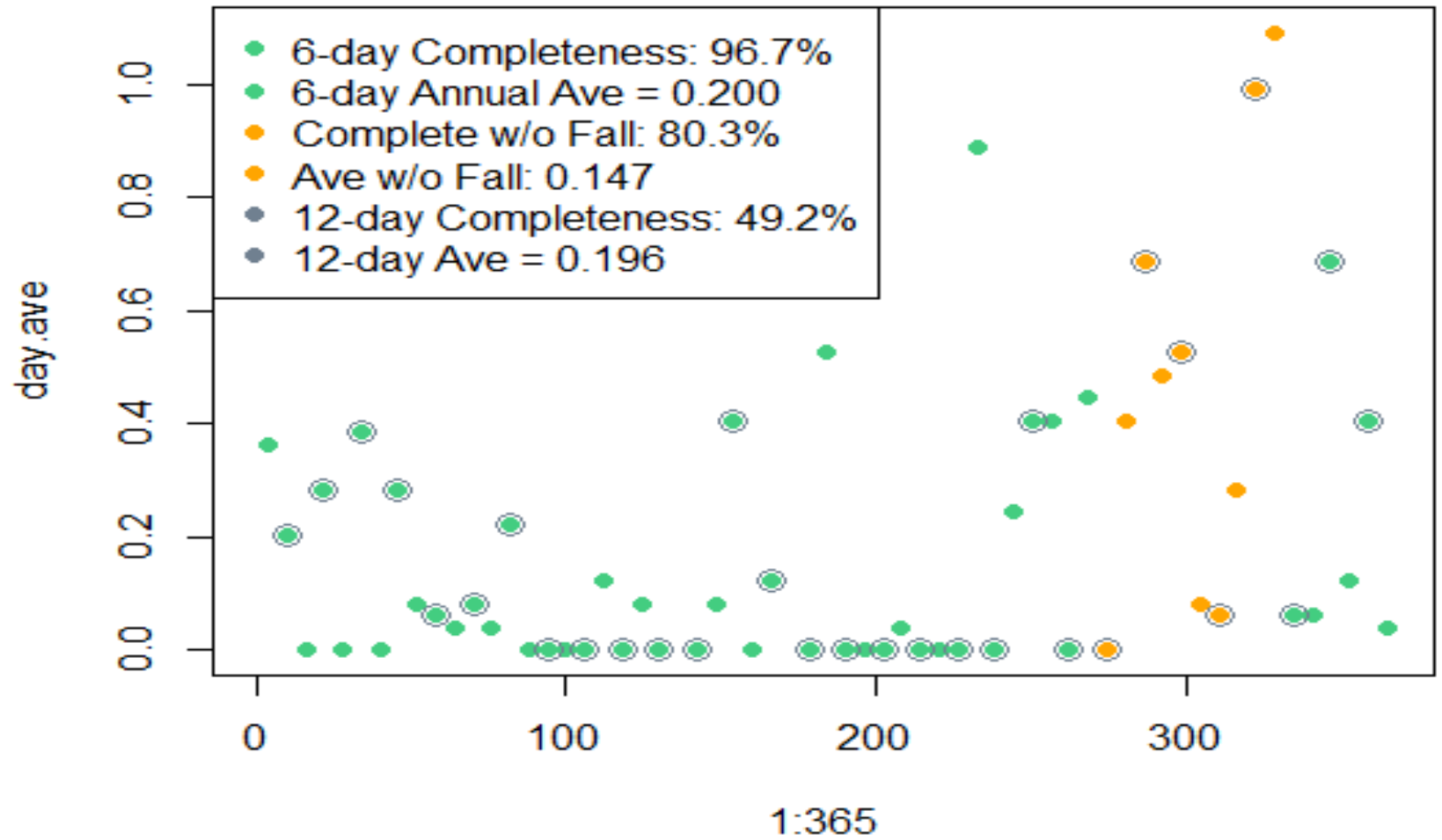

Daily Average of Ethylene dichloride

- 6-day Completeness: 96.7%
- 6-day Annual Ave = 0.200

day.ave

1:365

Ethylene Dichloride

Daily Average of Ethylene dichloride

Legend:
- 6-day Completeness: 96.7%
- 6-day Annual Ave = 0.200
- Completeness w/o Fall: 80.3%
- Ave w/o Fall: 0.147

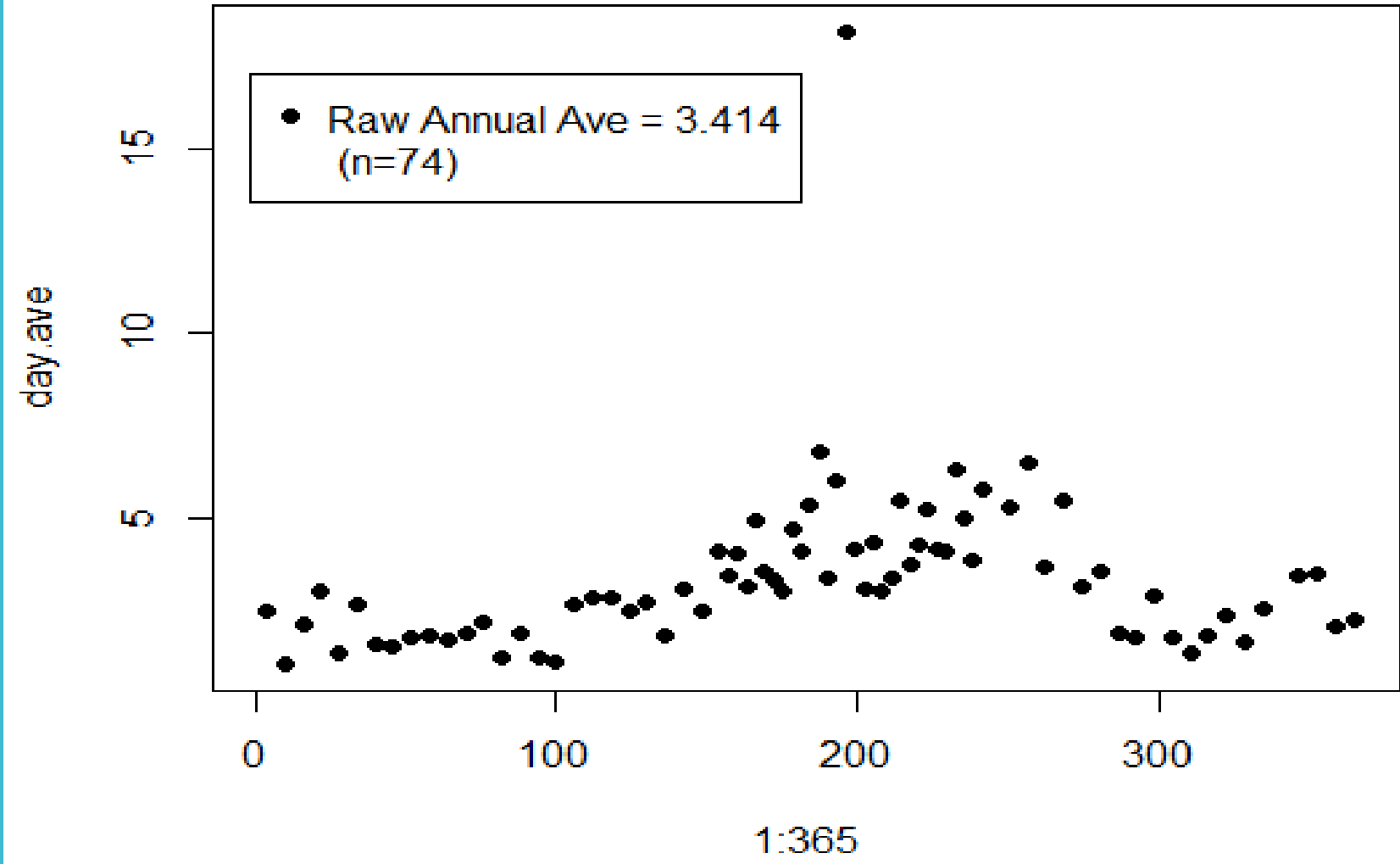y-axis: day.ave
x-axis: 1:365

Ethylene Dichloride

Formaldehyde

Complete over time on a 6-day scale

- Consider average across 6-day measurements
- Consider a >120% relative complete analysis: over-weighted seasonally due to extra summer measurements
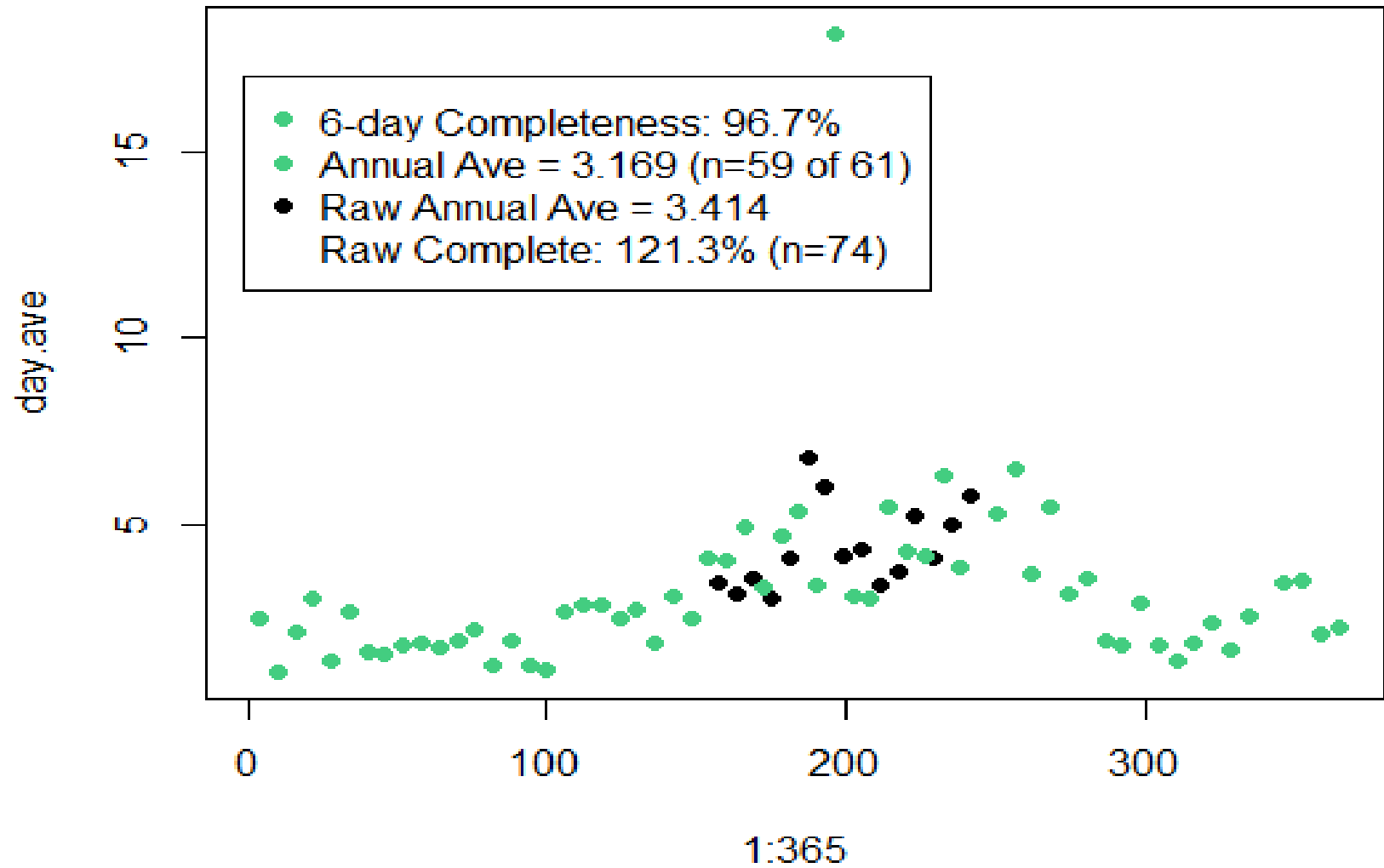- Compare effect on the averages
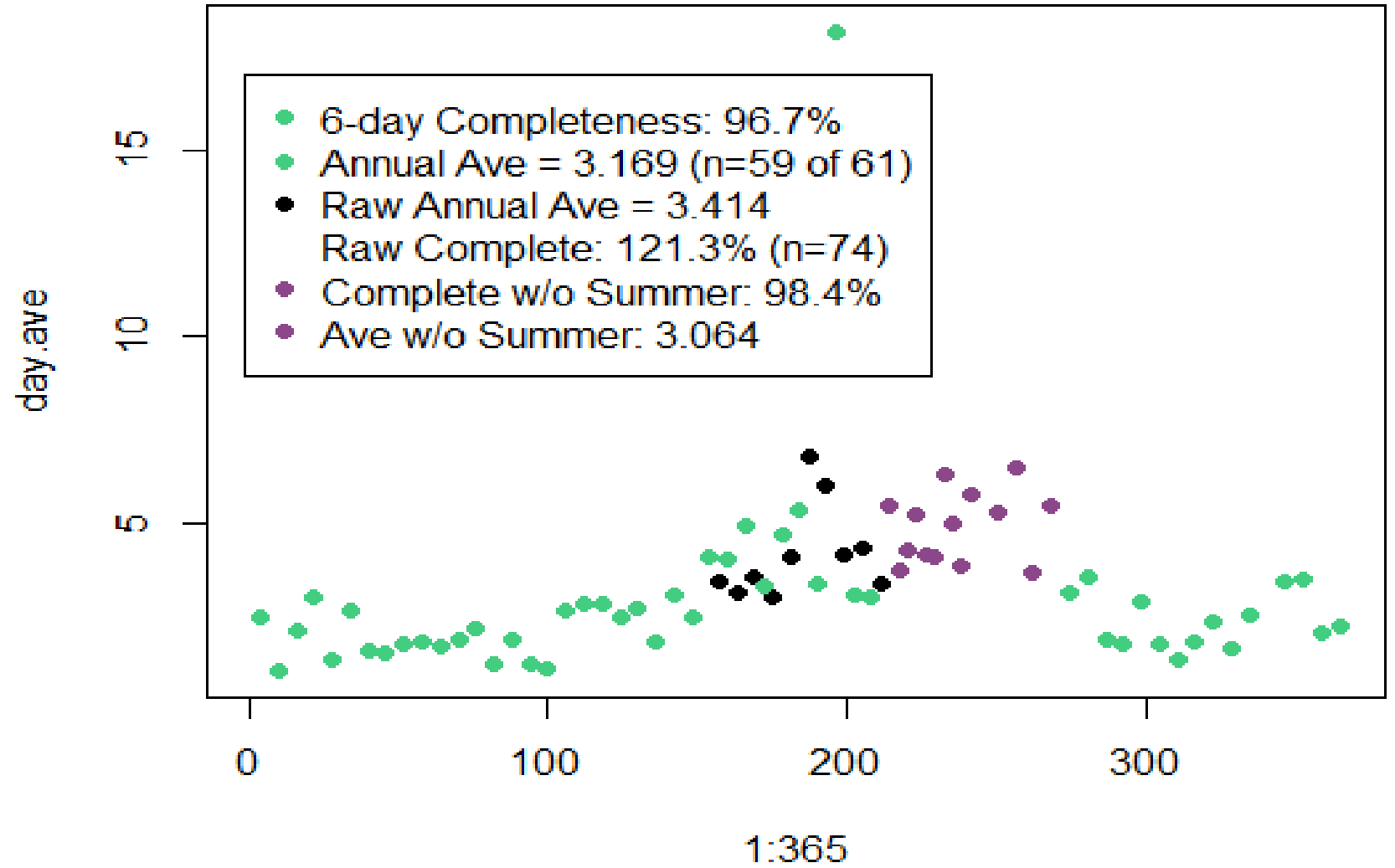
Formaldehyde

Daily Average of Formaldehyde

Formaldehyde

Daily Average of Formaldehyde

Formaldehyde

Take Away Message

- Completeness has a significant effect on trends analysis
- Caution must be used
- Seasonality is a big consideration and is tied into the completeness question.
- Seasonal effects can affect bias

Thank you