

Monthly Report (00)

201712 Data Set

Monday 15th January, 2018

Prepared for

Statistics for Physical and Engineering Sciences

by

Jamie Riggs, Ph.D.

Principal Statistician
Statistics for Physical and Engineering Sciences Institute

1 Introduction

The process of reporting monthly Sunspot numbers consists of submitting individual observer's daily counts for a specific month to the AAVSO Solar Section. These data are maintained in a SQL database. The monthly data then are extracted for analysis using the R statistics package (<http://www.R-project.org/>). This report is the portion of the analysis concerned with both the raw daily average counts and the data Accuracy, Consistency, and Completeness measures for a particular month. The checks are used to scrub or filter the data to assure only error-free data are used to determine the monthly sunspot number.

This report consists of four sections: the raw daily average counts (Section 2), the known data errors (Section 3), the processed counts using a Generalized Linear Mixed Model to produce the relative sunspot numbers (Section 4), and supporting information on the model construction (Section 5).

The raw daily average of counts consist of submitted counts from all observers who provided data in the particular month. These averaged counts are reported by the day of the month, and are either from data not scrubbed or corrected data. The table captions indicate which. The errors, if any, are reported according to type.

The Error Tables section contains reported errors on missing data, inconsistencies in year and month, inconsistencies in the reported day number (1-31), seeing coding errors, number of annual observations by observer, and inconsistencies between the reported Wolf number and the calculated Wolf number from the group counts and sunspot counts, among other errors that are given in that section.

The relative sunspot numbers R_a section contains the sunspot numbers after the submitted data are scrubbed and modeled by a Generalized Linear Mixed Model (GLMM). The GLMM is a statistical model that accounts for variation due to random effects and fixed effects. For the R_a model random effects include the AAVSO observer as these observers are a selection from all possible observers, and the fixed effects include seeing conditions at one of four possible levels. More details on GLMM are available in a paper (GLMM05) on the sunspot counts research page. The paper title is *A Generalized Linear Mixed Model for Enumerated Sunspots*.

The supporting information for the model is provided for clarification.

2 Raw Daily Average Counts

The reported raw daily average counts have been checked for errors and inconsistencies, and no known errors are present. All observers whose submissions qualify through this month's scrubbing process are represented in Figure 1 and Table 1.

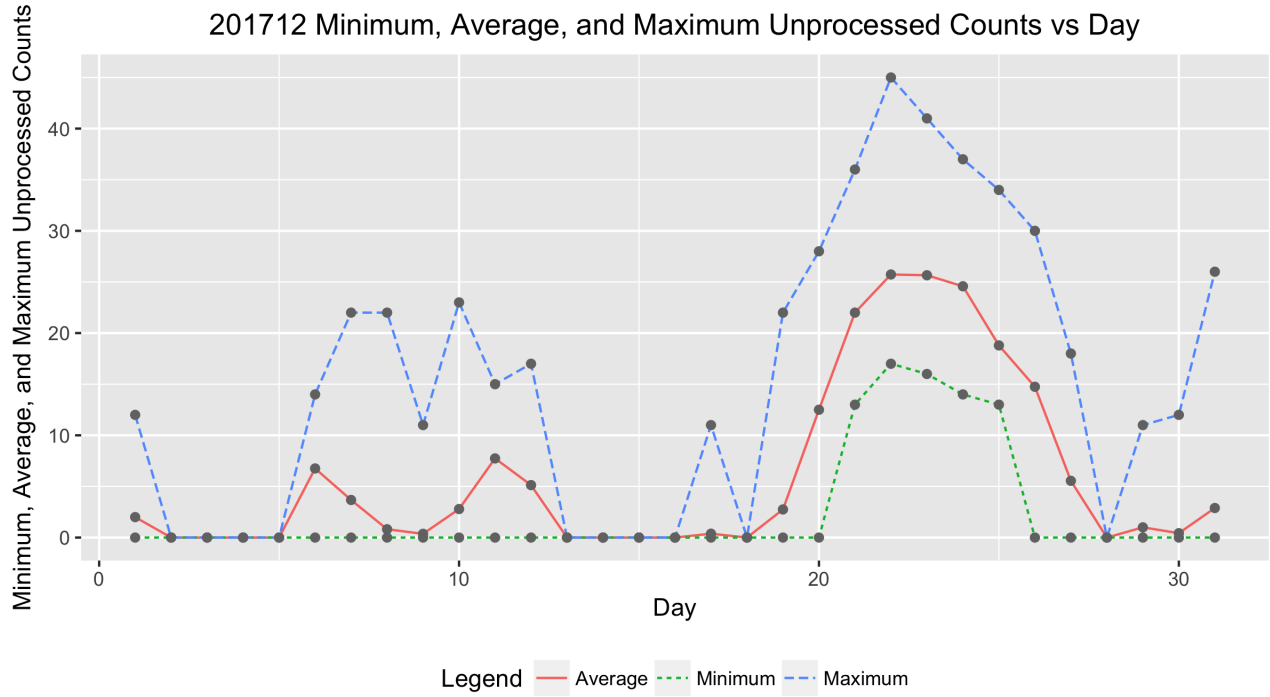


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201712 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	28.0000	0.0000	2.0000	12.0000
2.0000	31.0000	0.0000	0.0000	0.0000
3.0000	34.0000	0.0000	0.0000	0.0000
4.0000	28.0000	0.0000	0.0000	0.0000
5.0000	25.0000	0.0000	0.0000	0.0000
6.0000	29.0000	0.0000	6.7586	14.0000
7.0000	34.0000	0.0000	3.6765	22.0000
8.0000	27.0000	0.0000	0.8148	22.0000
9.0000	30.0000	0.0000	0.3667	11.0000
10.0000	33.0000	0.0000	2.7879	23.0000
11.0000	27.0000	0.0000	7.7407	15.0000
12.0000	31.0000	0.0000	5.1290	17.0000
13.0000	25.0000	0.0000	0.0000	0.0000
14.0000	29.0000	0.0000	0.0000	0.0000
15.0000	26.0000	0.0000	0.0000	0.0000
16.0000	38.0000	0.0000	0.0000	0.0000
17.0000	29.0000	0.0000	0.3793	11.0000
18.0000	27.0000	0.0000	0.0000	0.0000
19.0000	28.0000	0.0000	2.7500	22.0000
20.0000	28.0000	0.0000	12.5000	28.0000
21.0000	26.0000	13.0000	22.0000	36.0000
22.0000	26.0000	17.0000	25.7308	45.0000
23.0000	32.0000	16.0000	25.6562	41.0000
24.0000	26.0000	14.0000	24.5769	37.0000
25.0000	29.0000	13.0000	18.7931	34.0000
26.0000	28.0000	0.0000	14.7500	30.0000
27.0000	31.0000	0.0000	5.5484	18.0000
28.0000	31.0000	0.0000	0.0000	0.0000
29.0000	22.0000	0.0000	1.0000	11.0000
30.0000	28.0000	0.0000	0.4286	12.0000
31.0000	27.0000	0.0000	2.8889	26.0000

3 Error Tables

Data are for the month of December 2017. No errors were found, and hence no errors are reported.

4 Relative Sunspot Numbers

All data errors, if any, have been corrected prior to determining the following relative sunspot numbers. A Generalized Linear Mixed Model (GLMM) was constructed to provide monthly sunspot numbers (see Table 2). The GLMM treats observer as a random effect, with year, month, seeing conditions, observer rank, and dual submission to both AAVSO and SILSO as fixed effects.

Figure 2 shows the monthly R_a numbers for the years and months (ym) in Table 2. The solid cyan curve that connects the cyan X's are the GLMM model estimates given in 2. The dotted black curves on either side of the cyan curve depict a 99% confidence band about the GLMM estimates. The confidence band uses the large sample approximation based on the Gaussian distribution. The dashed red curve connecting the red O's are the SILSO values for the monthly sequence.

The tan box plots for each month are the actual observations submitted by the AAVSO observers. The heavy solid lines approximately midway in the boxes represent the count medians. The box of the box plot represents the InterQuartile Range (IQR), which depicts from the 25th through the 75th quartiles. The lower and upper whiskers extend 1.5 times the IQR below the 25th quartile, and 1.5 times the IQR above the 75th quartile. The black circles below and above the whiskers traditionally are considered outliers, but with GLMM modeling, they are observations that comprise overdispersion. Overdispersion skews the counts data from a true Poisson distribution. The GLMM adjusts for this overdispersion.

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2010.05	24.3651	23.9319	24.7982	8.4000	8.7000
2010.06	20.0609	19.7272	20.3945	11.0000	13.6000
2010.07	21.9001	21.5826	22.2175	15.2000	16.1000
2010.08	22.8658	22.4901	23.2415	18.3000	19.6000
2010.09	25.6626	25.2395	26.0856	22.8000	25.2000
2010.10	24.2019	23.7840	24.6199	21.0000	23.5000
2010.11	24.7028	24.2542	25.1513	20.9000	21.6000
2010.12	21.7528	21.3142	22.1913	13.9000	14.5000
2011.01	76.7685	75.1799	78.3571	17.7000	18.7000
2011.02	65.4824	64.0586	66.9063	29.1000	29.6000
2011.03	69.8969	68.5666	71.2272	48.0000	55.8000
2011.04	77.8346	76.4401	79.2292	47.3000	54.4000
2011.05	78.2342	76.9148	79.5535	37.3000	41.5000
2011.06	64.4033	63.2763	65.5302	35.2000	37.0000
2011.07	69.5530	68.3633	70.7427	41.5000	43.8000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2011.08	73.4860	72.3037	74.6683	42.4000	50.5000
2011.09	81.1134	79.7142	82.5126	73.8000	78.0000
2011.10	76.5463	75.2621	77.8304	78.9000	88.0000
2011.11	77.8680	76.2442	79.4918	84.6000	96.7000
2011.12	67.5220	66.1343	68.9096	65.8000	73.0000
2012.01	82.4270	80.8149	84.0390	55.8000	58.2000
2012.02	69.2606	67.8593	70.6620	29.2000	33.1000
2012.03	74.4409	73.1282	75.7535	53.1000	64.1000
2012.04	81.6885	80.2670	83.1101	51.4000	55.2000
2012.05	83.6234	82.2495	84.9973	61.8000	69.0000
2012.06	68.1779	67.0235	69.3324	59.7000	64.5000
2012.07	74.0133	72.8031	75.2236	64.2000	51.3000
2012.08	75.4064	74.1994	76.6134	57.7000	63.1000
2012.09	83.7607	82.3116	85.2098	57.7000	61.5000
2012.10	79.9247	78.4672	81.3821	48.3000	53.3000
2012.11	81.2968	79.6721	82.9215	56.7000	61.4000
2012.12	70.6451	69.1175	72.1727	37.4000	40.8000
2013.01	91.6463	89.9028	93.3898	63.8000	62.9000
2013.02	77.1497	75.6039	78.6955	37.8000	38.0000
2013.03	80.2922	78.6687	81.9158	50.6000	57.9000
2013.04	89.0981	87.5409	90.6553	70.6000	72.4000
2013.05	89.1710	87.5922	90.7498	77.4000	78.7000
2013.06	74.1402	72.8358	75.4447	51.0000	52.5000
2013.07	79.4997	78.2291	80.7703	57.0000	57.0000
2013.08	82.5579	81.2364	83.8794	60.0000	66.0000
2013.09	90.4192	88.8148	92.0236	34.6000	36.9000
2013.10	85.1792	83.6090	86.7495	74.5000	85.6000
2013.11	85.0895	83.1923	86.9867	73.9000	77.6000
2013.12	76.0462	74.4317	77.6606	77.8000	90.3000
2014.01	106.7630	104.5292	108.9969	77.4000	82.0000
2014.02	91.6628	89.8664	93.4591	93.9000	102.8000
2014.03	97.6575	95.8890	99.4260	80.9000	92.2000
2014.04	108.5154	106.6326	110.3983	76.9000	84.7000
2014.05	109.2722	107.4176	111.1268	72.3000	75.2000
2014.06	90.6598	89.1337	92.1859	67.2000	71.0000
2014.07	96.9384	95.3329	98.5439	72.5000	72.5000
2014.08	100.7997	99.2321	102.3672	71.2000	74.7000
2014.09	111.5332	109.5767	113.4897	83.2000	87.6000
2014.10	104.7886	102.8677	106.7096	59.5000	60.6000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2014.11	105.6872	103.4993	107.8752	65.8000	71.1000
2014.12	92.5246	90.3908	94.6583	75.8000	78.0000
2015.01	66.0628	64.7467	67.3790	65.9000	67.0000
2015.02	55.3009	54.0881	56.5137	42.4000	44.8000
2015.03	59.6663	58.5779	60.7546	38.0000	38.4000
2015.04	65.8800	64.7005	67.0594	49.0000	54.4000
2015.05	66.6321	65.5314	67.7328	56.3000	58.8000
2015.06	55.2063	54.2115	56.2012	50.2000	68.3000
2015.07	58.4827	57.4857	59.4796	47.9000	65.8000
2015.08	62.0897	61.0472	63.1321	39.5000	57.2000
2015.09	67.7929	66.5634	69.0223	49.2000	72.1000
2015.10	64.1893	62.9521	65.4266	39.3000	48.3000
2015.11	65.4696	64.0527	66.8865	39.6000	55.9000
2015.12	57.3821	56.1250	58.6391	36.4000	44.8000
2016.01	36.1470	35.3954	36.8986	33.7000	43.3000
2016.02	30.4107	29.7793	31.0422	38.3000	46.8000
2016.03	32.2989	31.6570	32.9408	30.5000	38.9000
2016.04	35.5253	34.8496	36.2010	26.6000	30.9000
2016.05	36.0311	35.3729	36.6892	33.7000	48.4000
2016.06	29.5311	29.0282	30.0339	13.1000	19.5000
2016.07	31.8783	31.3657	32.3910	21.2000	27.5000
2016.08	33.4626	32.8745	34.0506	33.0000	47.9000
2016.09	37.3989	36.7167	38.0812	27.7000	37.1000
2016.10	35.0189	34.3446	35.6932	22.7000	31.7000
2016.11	35.3058	34.5693	36.0424	14.0000	22.2000
2016.12	31.3500	30.6821	32.0178	11.1000	20.0000
2017.01	19.6894	19.2761	20.1027	18.4000	26.2000
2017.02	16.6348	16.2705	16.9990	14.4000	20.6000
2017.03	17.8211	17.4829	18.1594	11.3000	15.5000
2017.04	19.8251	19.4760	20.1743	21.6000	33.2000
2017.05	19.7813	19.4403	20.1222	12.5000	18.1000
2017.06	16.2288	15.9588	16.4989	15.5000	19.3000
2017.07	17.5935	17.3147	17.8724	11.5000	16.3000
2017.08	18.3915	18.0788	18.7043	22.8000	35.7000
2017.09	20.7554	20.3726	21.1383	34.6000	42.9000
2017.10	19.0570	18.6830	19.4309	10.5000	11.0000
2017.11	19.0532	18.6514	19.4550	4.2000	5.6000
2017.12	16.7713	16.5095	17.0331	4.0000	4.6000

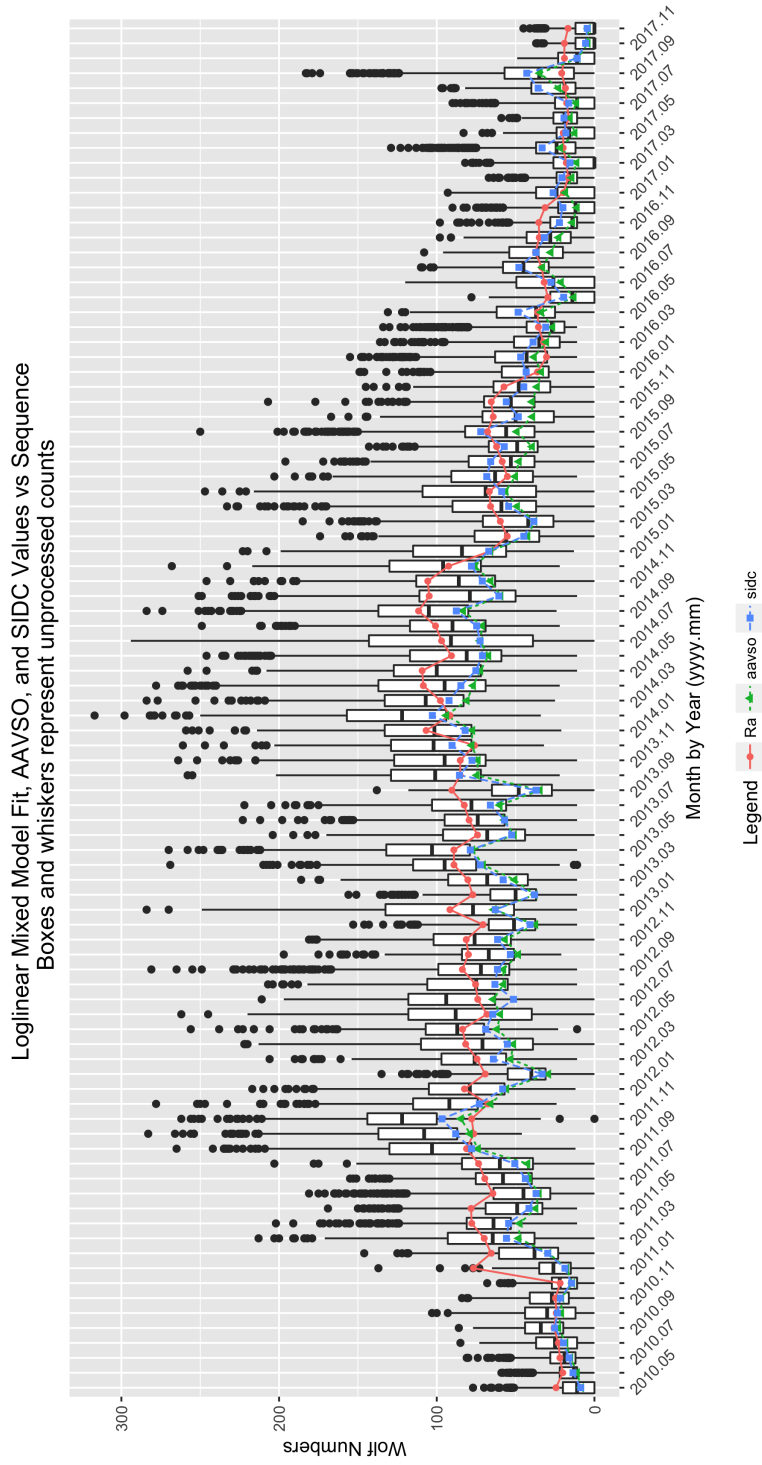


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

The GLMM parameter estimates and measures of importance in the determining the monthly R_a values are given in Table 3. The parameter estimates and levels of statistical significance are determined for the residual error size combined with the observer random effect error size. Thus, the parameter estimates are adjusted for the random effect of observer. The significance level is set at 0.05. Any $\Pr(>|z|)$ values equal to or less than 0.05 are considered statistically significant.

Table 3: 201712 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2872	0.0327	100.4674	0.0000
seeF	-0.2183	0.0061	-35.7132	0.0000
seeG	-0.1153	0.0053	-21.6684	0.0000
seeM	-0.1951	0.0250	-7.8185	0.0000
seeP	-0.3265	0.0088	-37.2767	0.0000
sidc1	0.1114	0.0722	1.5424	0.1230
year2011	1.1158	0.0119	93.5906	0.0000
year2012	1.1532	0.0119	97.2152	0.0000
year2013	1.2492	0.0118	105.5682	0.0000
year2014	1.4462	0.0117	123.2909	0.0000
year2015	0.9614	0.0121	79.2716	0.0000
year2016	0.3445	0.0130	26.5753	0.0000
year2017	-0.2587	0.0141	-18.3947	0.0000
mon2	-0.1632	0.0097	-16.8569	0.0000
mon3	-0.1077	0.0090	-11.9104	0.0000
mon4	-0.0133	0.0087	-1.5253	0.1272
mon5	-0.0122	0.0086	-1.4222	0.1550
mon6	-0.2090	0.0090	-23.1674	0.0000
mon7	-0.1406	0.0087	-16.1499	0.0000
mon8	-0.0930	0.0085	-10.8831	0.0000
mon9	0.0163	0.0086	1.9008	0.0573
mon10	-0.0401	0.0088	-4.5536	0.0000
mon11	-0.0168	0.0092	-1.8235	0.0682
mon12	-0.1397	0.0094	-14.8478	0.0000

The year effect levels are given as year2011, year2012, and year2013. The yearly effect is significant as $\Pr(> |z|) < 0.05$. So the year in which the observations are made is commensurate with the expected rise toward and anticipated sunspot number maximum. Similarly, the monthly effect, denoted as mon2 through mon12, is significant at the 0.05 level.

The seeing conditions account for a significant amount of deviation in sunspot numbers. The seeing conditions are denoted as seeF (Fair), seeG (Good), and seeP (Poor), and are significant at the 0.05 level. Therefore, seeing conditions influence the reported sunspot numbers, as intuition anticipates.

The level of observer experience (denoted r1000B through r5000H, which is least to most expe-

ience) is not significant at the 0.05 significance level. It therefore does not contribute to changes in the monthly sunspot numbers.

Whether an observer contributes counts to the SILSO as well as the AAVSO (silsoy) is not significant at the 0.05 level, and hence we conclude that those observers who contribution to both institutions tend to differ from those observers contributing only to the AAVSO.

5 Supporting Information

Table 4: 201712 Summary of Sunspot Numbers

jd	year	mon	day	obs
Min. :2455318	Min. :2010	Min. : 1.000	Min. : 1.00	ARAG : 2739
1st Qu.:2456108	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00	BRAB : 2555
Median :2456775	Median :2014	Median : 7.000	Median :16.00	CHAG : 2533
Mean :2456768	Mean :2014	Mean : 6.711	Mean :15.72	MJAF : 2338
3rd Qu.:2457459	3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.00	KNJS : 2266
Max. :2458119	Max. :2017	Max. :12.000	Max. :31.00	BROB : 2252 (Other):77659

Table 5: Summary of Sunspot Numbers

g	s	w
Min. : 0.000	Min. : 0.00	Min. : 0.00
1st Qu.: 2.000	1st Qu.: 6.00	1st Qu.: 25.00
Median : 3.000	Median : 16.00	Median : 53.00
Mean : 3.817	Mean : 22.88	Mean : 61.05
3rd Qu.: 6.000	3rd Qu.: 33.00	3rd Qu.: 89.00
Max. :19.000	Max. :204.00	Max. :317.00

Table 6: Summary of Sunspot Numbers

aperture	eyep	foclen	mag	k	ow
Min. : 0.0	Min. : 0.00	Min. : 0	Min. : 0.0	Min. :0.0000	Min. :0.0000
1st Qu.: 76.0	1st Qu.: 2.00	1st Qu.: 800	1st Qu.: 40.0	1st Qu.:0.0000	1st Qu.:0.0000
Median : 90.0	Median : 13.00	Median :1000	Median : 60.0	Median :0.5810	Median :1.0000
Mean : 110.7	Mean : 15.34	Mean :1116	Mean : 187.4	Mean :0.4043	Mean :0.5658
3rd Qu.: 125.0	3rd Qu.: 23.00	3rd Qu.:1296	3rd Qu.: 77.0	3rd Qu.:0.6910	3rd Qu.:1.0000
Max. :1524.0	Max. :2010.00	Max. :4300	Max. :4591.0	Max. :1.1200	Max. :1.0000

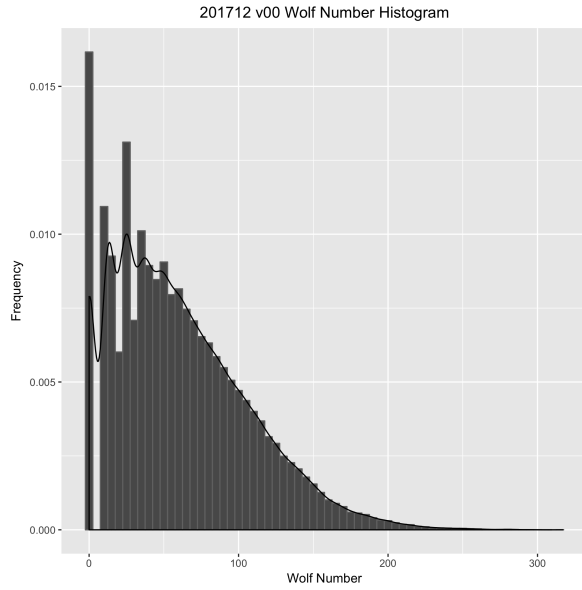


Figure 3: Box plots of raw Wolf number (w) by observer rank.

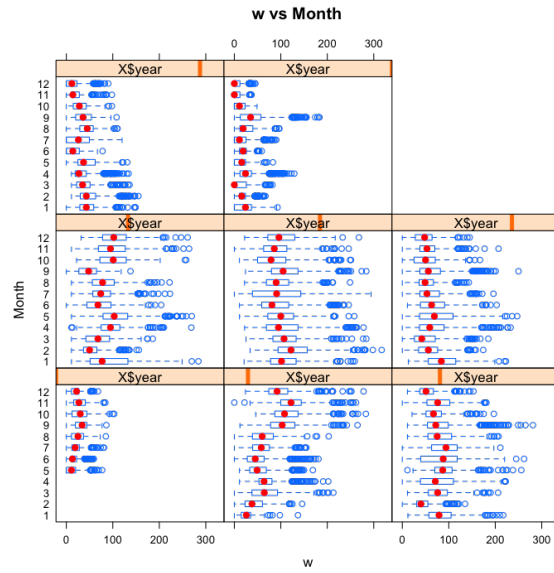


Figure 4: Box plots of raw Wolf number (w) by month and year.

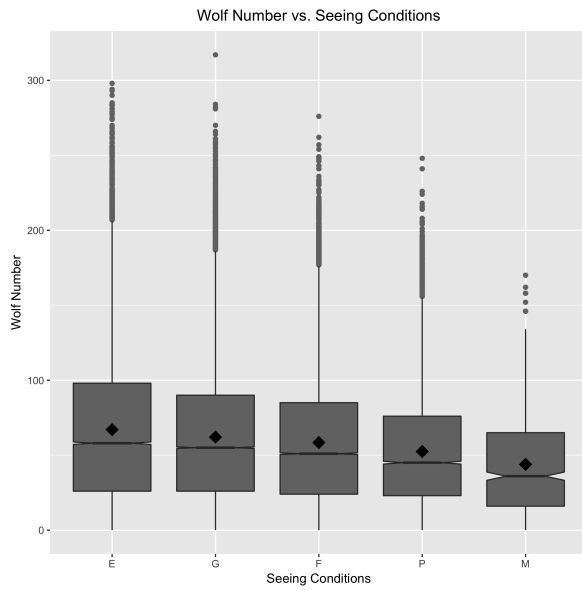


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

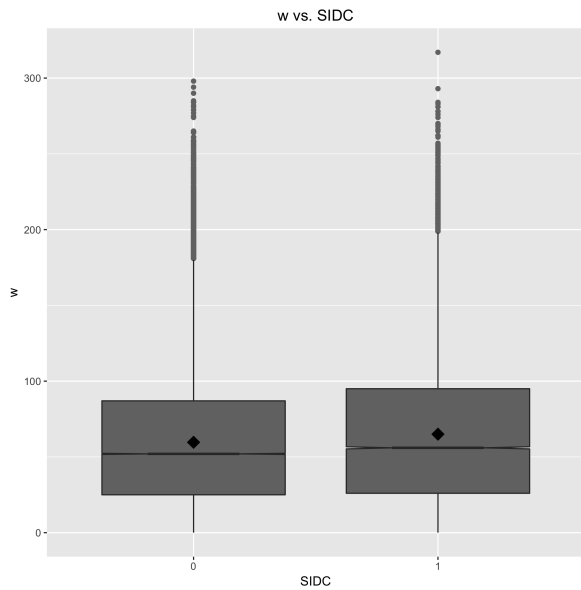


Figure 6: Box plots of raw Wolf number (w) by organization.

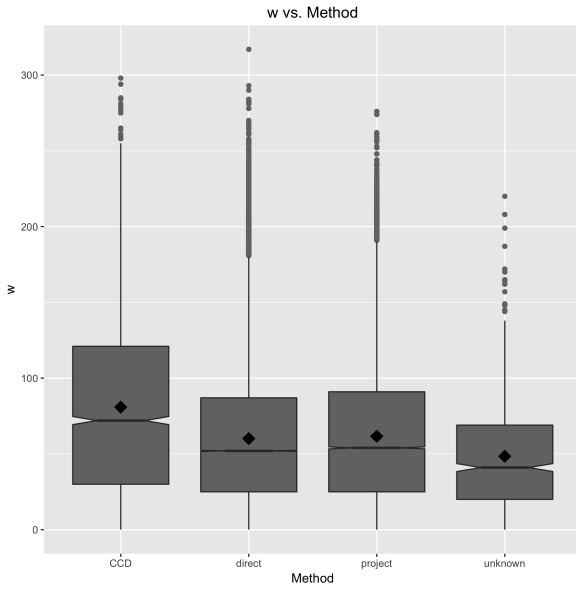


Figure 7: Box plots of raw Wolf number (w) by observer rank.

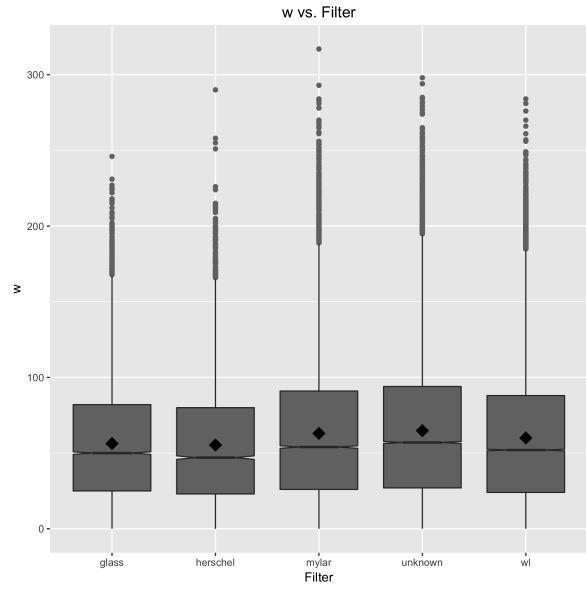


Figure 8: Box plots of raw Wolf number (w) by month and year.

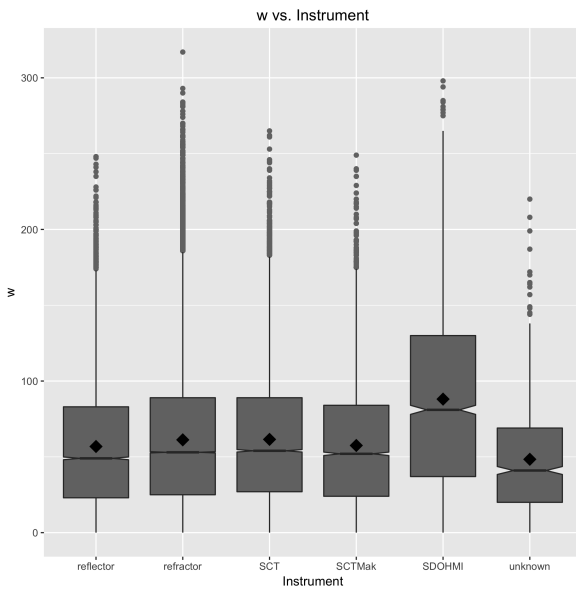


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

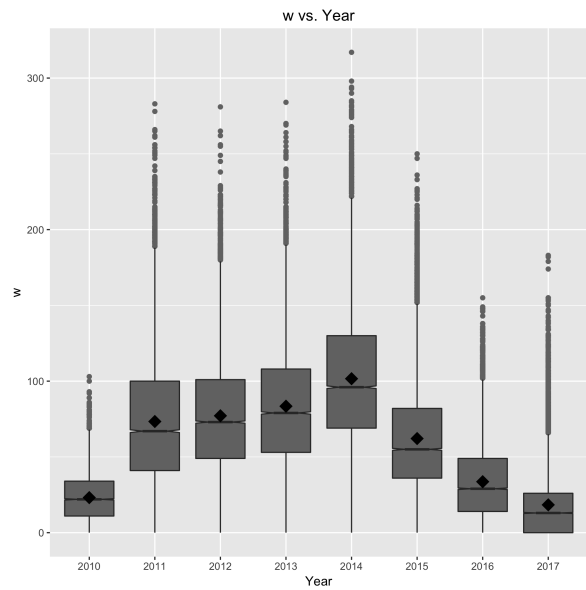


Figure 10: Box plots of raw Wolf number (w) by organization.

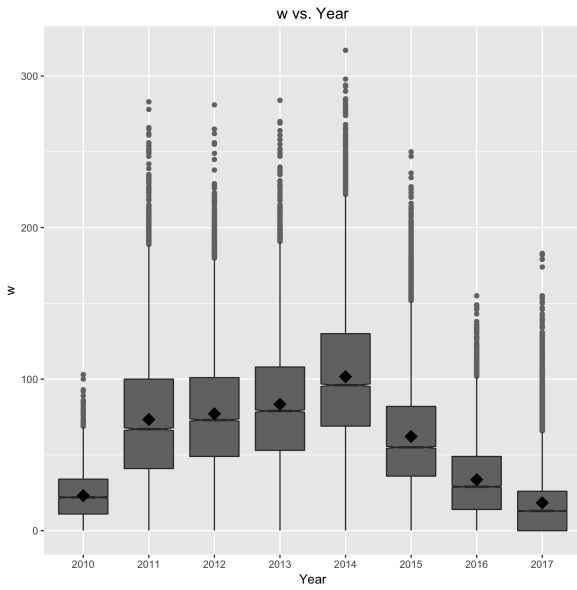


Figure 11: Box plots of raw Wolf number (w) by year.

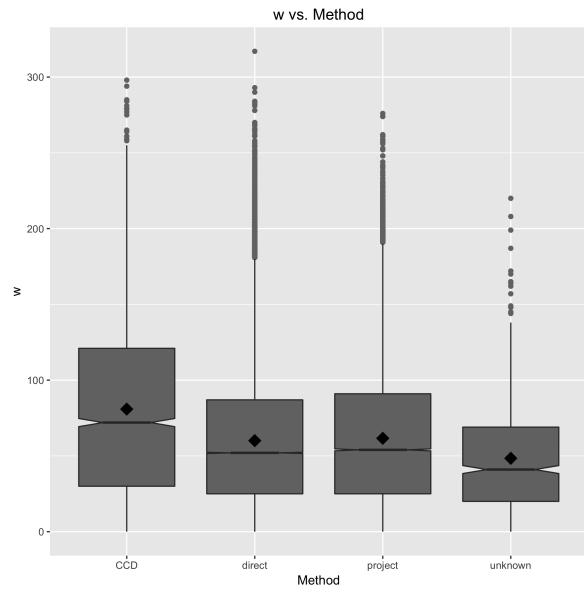


Figure 12: Box plots of raw Wolf number (w) by observing method.