Monthly Report (00)

201801 Data Set

Sunday 11th February, 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.

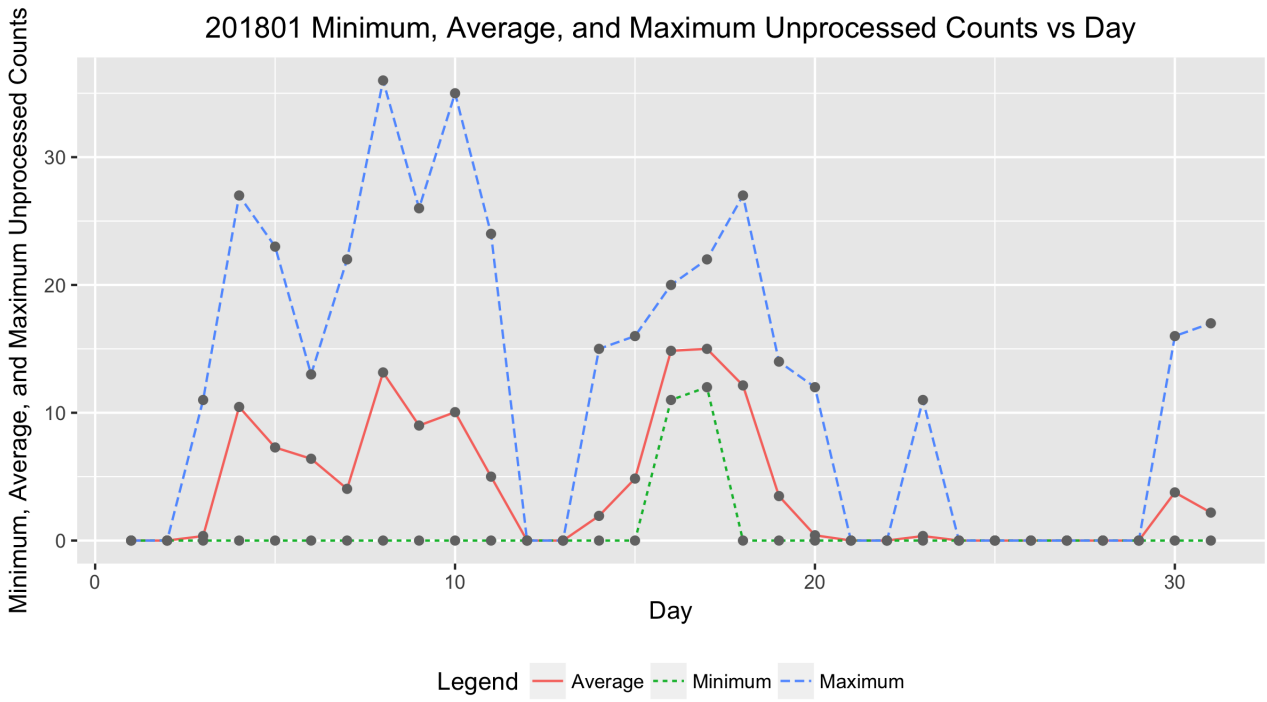

Figure 1: Raw average sunspot count by day of the month.

Table 1: 201801 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	33.0000	0.0000	0.0000	0.0000
2.0000	30.0000	0.0000	0.0000	0.0000
3.0000	31.0000	0.0000	0.3548	11.0000
4.0000	33.0000	0.0000	10.4545	27.0000
5.0000	28.0000	0.0000	7.2857	23.0000
6.0000	35.0000	0.0000	6.4000	13.0000
7.0000	22.0000	0.0000	4.0455	22.0000
8.0000	13.0000	0.0000	13.1538	36.0000
9.0000	17.0000	0.0000	9.0000	26.0000
10.0000	21.0000	0.0000	10.0476	35.0000
11.0000	19.0000	0.0000	5.0000	24.0000
12.0000	23.0000	0.0000	0.0000	0.0000
13.0000	30.0000	0.0000	0.0000	0.0000
14.0000	29.0000	0.0000	1.9310	15.0000
15.0000	20.0000	0.0000	4.8500	16.0000
16.0000	26.0000	11.0000	14.8462	20.0000
17.0000	23.0000	12.0000	15.0000	22.0000
18.0000	29.0000	0.0000	12.1379	27.0000
19.0000	25.0000	0.0000	3.4800	14.0000
20.0000	29.0000	0.0000	0.4138	12.0000
21.0000	32.0000	0.0000	0.0000	0.0000
22.0000	22.0000	0.0000	0.0000	0.0000
23.0000	31.0000	0.0000	0.3548	11.0000
24.0000	27.0000	0.0000	0.0000	0.0000
25.0000	26.0000	0.0000	0.0000	0.0000
26.0000	35.0000	0.0000	0.0000	0.0000
27.0000	27.0000	0.0000	0.0000	0.0000
28.0000	29.0000	0.0000	0.0000	0.0000
29.0000	23.0000	0.0000	0.0000	0.0000
30.0000	30.0000	0.0000	3.7667	16.0000
31.0000	32.0000	0.0000	2.1875	17.0000

3 Error Tables

Data are for the month of January 2018. 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 25^{th} through the 75^{th} quartiles. The lower and upper whiskers extend 1.5 times the IQR below the 25^{th} quartile, and 1.5 times the IQR above the 75^{th} 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	24.3639	23.9304	24.7974	8.4000	8.7000
2010.06	20.0605	19.7268	20.3942	11.0000	13.6000
2010.07	21.8997	21.5822	22.2172	15.2000	16.1000
2010.08	22.8663	22.4905	23.2422	18.3000	19.6000
2010.09	25.6630	25.2397	26.0862	22.8000	25.2000
2010.10	24.2025	23.7844	24.6206	21.0000	23.5000
2010.11	24.7032	24.2544	25.1519	20.9000	21.6000
2010.12	21.7532	21.3144	22.1921	13.9000	14.5000
2011.01	76.7689	75.1793	78.3585	17.7000	18.7000
2011.02	65.4817	64.0569	66.9064	29.1000	29.6000
2011.03	69.8971	68.5661	71.2282	48.0000	55.8000
2011.04	77.8352	76.4399	79.2304	47.3000	54.4000
2011.05	78.2343	76.9143	79.5543	37.3000	41.5000
2011.06	64.4030	63.2756	65.5304	35.2000	37.0000
2011.07	69.5525	68.3624	70.7426	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	72.3033	74.6688	42.4000	50.5000
2011.09	81.1145	79.7147	82.5143	73.8000	78.0000
2011.10	76.5471	75.2623	77.8320	78.9000	88.0000
2011.11	77.8672	76.2426	79.4919	84.6000	96.7000
2011.12	67.5204	66.1320	68.9088	65.8000	73.0000
2012.01	82.4274	80.8139	84.0409	55.8000	58.2000
2012.02	69.2608	67.8582	70.6633	29.2000	33.1000
2012.03	74.4400	73.1265	75.7535	53.1000	64.1000
2012.04	81.6878	80.2652	83.1104	51.4000	55.2000
2012.05	83.6238	82.2489	84.9986	61.8000	69.0000
2012.06	68.1776	67.0224	69.3328	59.7000	64.5000
2012.07	74.0139	72.8029	75.2249	64.2000	51.3000
2012.08	75.4057	74.1979	76.6135	57.7000	63.1000
2012.09	83.7611	82.3111	85.2112	57.7000	61.5000
2012.10	79.9258	78.4674	81.3843	48.3000	53.3000
2012.11	81.2961	79.6701	82.9220	56.7000	61.4000
2012.12	70.6457	69.1168	72.1745	37.4000	40.8000
2013.01	91.6464	89.9015	93.3914	63.8000	62.9000
2013.02	77.1500	75.6029	78.6970	37.8000	38.0000
2013.03	80.2926	78.6680	81.9173	50.6000	57.9000
2013.04	89.0967	87.5384	90.6549	70.6000	72.4000
2013.05	89.1701	87.5902	90.7499	77.4000	78.7000
2013.06	74.1400	72.8348	75.4453	51.0000	52.5000
2013.07	79.4998	78.2284	80.7712	57.0000	57.0000
2013.08	82.5586	81.2363	83.8809	60.0000	66.0000
2013.09	90.4201	88.8147	92.0256	34.6000	36.9000
2013.10	85.1785	83.6073	86.7498	74.5000	85.6000
2013.11	85.0905	83.1919	86.9891	73.9000	77.6000
2013.12	76.0462	74.4306	77.6618	77.8000	90.3000
2014.01	106.7634	104.5279	108.9989	77.4000	82.0000
2014.02	91.6615	89.8640	93.4591	93.9000	102.8000
2014.03	97.6574	95.8878	99.4270	80.9000	92.2000
2014.04	108.5156	106.6315	110.3997	76.9000	84.7000
2014.05	109.2722	107.4165	111.1279	72.3000	75.2000
2014.06	90.6601	89.1331	92.1871	67.2000	71.0000
2014.07	96.9381	95.3317	98.5445	72.5000	72.5000
2014.08	100.7998	99.2314	102.3683	71.2000	74.7000
2014.09	111.5337	109.5760	113.4914	83.2000	87.6000
2014.10	104.7888	102.8666	106.7110	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	105.6862	103.4964	107.8760	65.8000	71.1000
2014.12	92.5252	90.3895	94.6609	75.8000	78.0000
2015.01	66.0603	64.7431	67.3775	65.9000	67.0000
2015.02	55.3010	54.0873	56.5147	42.4000	44.8000
2015.03	59.6663	58.5772	60.7554	38.0000	38.4000
2015.04	65.8800	64.6999	67.0601	49.0000	54.4000
2015.05	66.6318	65.5303	67.7333	56.3000	58.8000
2015.06	55.2068	54.2113	56.2023	50.2000	68.3000
2015.07	58.4825	57.4848	59.4801	47.9000	65.8000
2015.08	62.0881	61.0451	63.1312	39.5000	57.2000
2015.09	67.7932	66.5629	69.0235	49.2000	72.1000
2015.10	64.1906	62.9525	65.4288	39.3000	48.3000
2015.11	65.4718	64.0537	66.8898	39.6000	55.9000
2015.12	57.3828	56.1246	58.6410	36.4000	44.8000
2016.01	36.1469	35.3946	36.8992	33.7000	43.3000
2016.02	30.4110	29.7790	31.0430	38.3000	46.8000
2016.03	32.2985	31.6561	32.9410	30.5000	38.9000
2016.04	35.5261	34.8499	36.2024	26.6000	30.9000
2016.05	36.0321	35.3734	36.6907	33.7000	48.4000
2016.06	29.5302	29.0269	30.0334	13.1000	19.5000
2016.07	31.8774	31.3645	32.3904	21.2000	27.5000
2016.08	33.4623	32.8739	34.0507	33.0000	47.9000
2016.09	37.3998	36.7170	38.0825	27.7000	37.1000
2016.10	35.0191	34.3443	35.6939	22.7000	31.7000
2016.11	35.3066	34.5695	36.0437	14.0000	22.2000
2016.12	31.3490	30.6807	32.0173	11.1000	20.0000
2017.01	19.6909	19.2773	20.1046	18.4000	26.2000
2017.02	16.6360	16.2715	17.0005	14.4000	20.6000
2017.03	17.8220	17.4834	18.1605	11.3000	15.5000
2017.04	19.8260	19.4766	20.1755	21.6000	33.2000
2017.05	19.7821	19.4409	20.1232	12.5000	18.1000
2017.06	16.2299	15.9597	16.5001	15.5000	19.3000
2017.07	17.5950	17.3161	17.8740	11.5000	16.3000
2017.08	18.3928	18.0800	18.7056	22.8000	35.7000
2017.09	20.7512	20.3698	21.1326	34.6000	42.9000
2017.10	19.0537	18.6795	19.4279	10.5000	11.0000
2017.11	19.0516	18.6491	19.4541	4.2000	5.6000
2017.12	16.7713	16.5091	17.0335	4.0000	4.6000

Continued on next page

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

ym	Ra	lci99	uci99	aavso	sidc
2018.01	3.6530	3.5757	3.7303	3.1000	6.3000

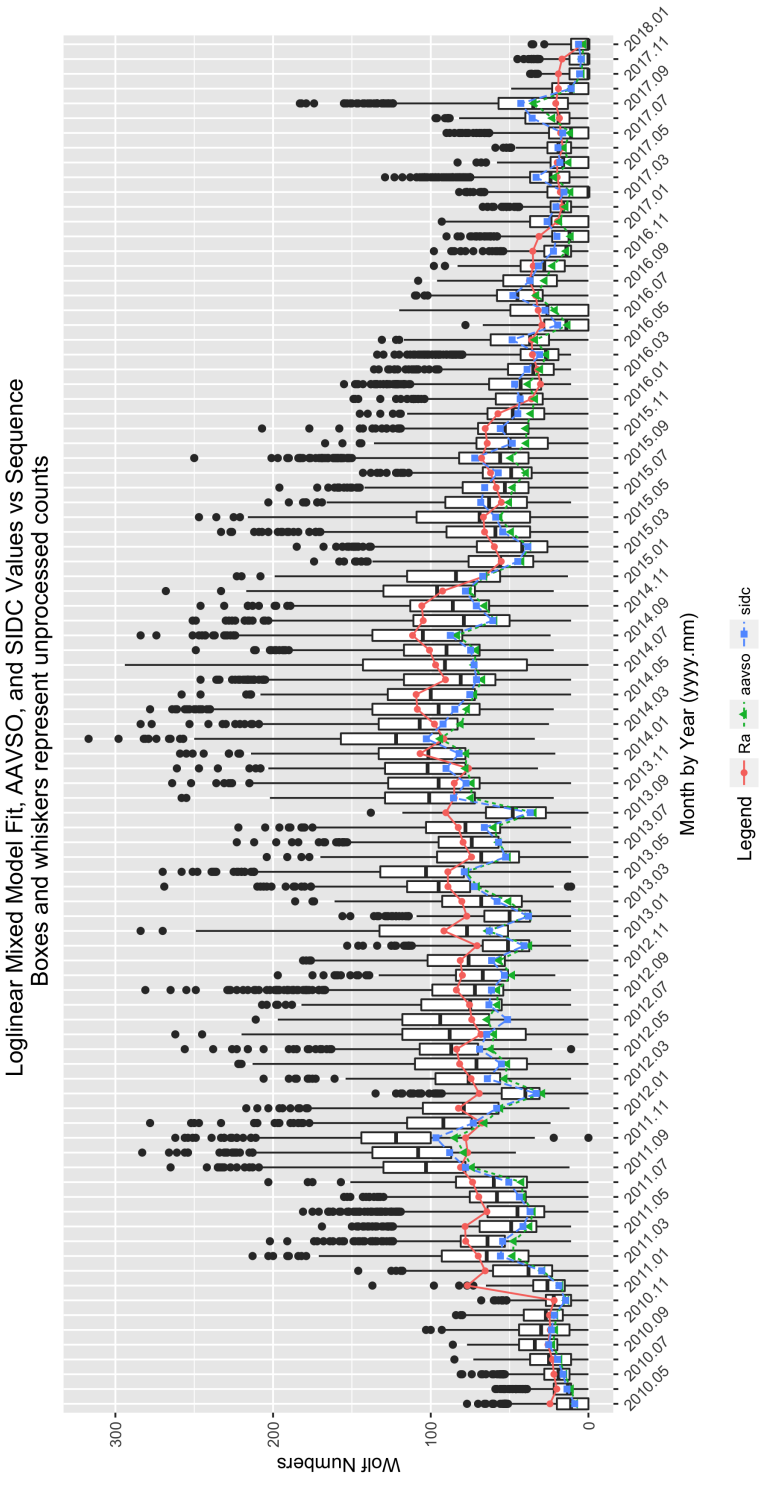

Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201801 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.2858	0.0327	100.5456	0.0000
seeF	-0.2183	0.0061	-35.7967	0.0000
seeG	-0.1153	0.0053	-21.7071	0.0000
seeM	-0.1951	0.0249	-7.8336	0.0000
seeP	-0.3264	0.0087	-37.3531	0.0000
sidc1	0.1090	0.0720	1.5137	0.1301
year2011	1.1158	0.0119	93.7745	0.0000
year2012	1.1531	0.0118	97.4047	0.0000
year2013	1.2491	0.0118	105.7749	0.0000
year2014	1.4461	0.0117	123.5330	0.0000
year2015	0.9614	0.0121	79.4286	0.0000
year2016	0.3444	0.0129	26.6249	0.0000
year2017	-0.2588	0.0140	-18.4321	0.0000
year2018	-1.9508	0.0765	-25.4975	0.0000
mon2	-0.1632	0.0097	-16.8908	0.0000
mon3	-0.1077	0.0090	-11.9341	0.0000
mon4	-0.0133	0.0087	-1.5269	0.1268
mon5	-0.0122	0.0086	-1.4244	0.1543
mon6	-0.2090	0.0090	-23.2125	0.0000
mon7	-0.1406	0.0087	-16.1806	0.0000
mon8	-0.0930	0.0085	-10.9025	0.0000
mon9	0.0163	0.0085	1.9078	0.0564
mon10	-0.0401	0.0088	-4.5609	0.0000
mon11	-0.0168	0.0092	-1.8264	0.0678
mon12	-0.1397	0.0094	-14.8784	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 experience) 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: 201801 Summary of Sunspot Numbers

jd	year	mon	day	obs
Min. :2455318	Min. :2010	Min. : 1.00	Min. : 1.00	ARAG : 2769
1st Qu.:2456113	1st Qu.:2012	1st Qu.: 4.00	1st Qu.: 8.00	CHAG : 2556
Median :2456786	Median :2014	Median : 7.00	Median :16.00	BRAB : 2555
Mean :2456780	Mean :2014	Mean : 6.66	Mean :15.72	MJAF : 2369
3rd Qu.:2457478	3rd Qu.:2016	3rd Qu.: 9.00	3rd Qu.:23.00	KNJS : 2297
Max. :2458150	Max. :2018	Max. :12.00	Max. :31.00	BROB : 2276
				(Other):78350

Table 5: Summary of Sunspot Numbers

g	s	w	aperture
Min. : 0.000	Min. : 0.00	Min. : 0.00	Min. : 0.0
1st Qu.: 2.000	1st Qu.: 6.00	1st Qu.: 25.00	1st Qu.: 76.0
Median : 3.000	Median : 16.00	Median : 52.00	Median : 90.0
Mean : 3.786	Mean : 22.69	Mean : 60.54	Mean : 110.7
3rd Qu.: 5.000	3rd Qu.: 33.00	3rd Qu.: 88.00	3rd Qu.: 125.0
Max. :19.000	Max. :204.00	Max. :317.00	Max. :1524.0

Table 6: Summary of Sunspot Numbers

eyep	foclen	mag	k	ow
Min. : 0.00	Min. : 0	Min. : 0.0	Min. :0.000	Min. :0.0000
1st Qu.: 2.00	1st Qu.: 800	1st Qu.: 40.0	1st Qu.:0.000	1st Qu.:0.0000
Median : 13.00	Median :1000	Median : 60.0	Median :0.568	Median :1.0000
Mean : 15.98	Mean :1114	Mean : 187.6	Mean :0.404	Mean :0.5654
3rd Qu.: 23.00	3rd Qu.:1296	3rd Qu.: 76.0	3rd Qu.:0.691	3rd Qu.:1.0000
Max. :2010.00	Max. :4300	Max. :4591.0	Max. :1.120	Max. :1.0000

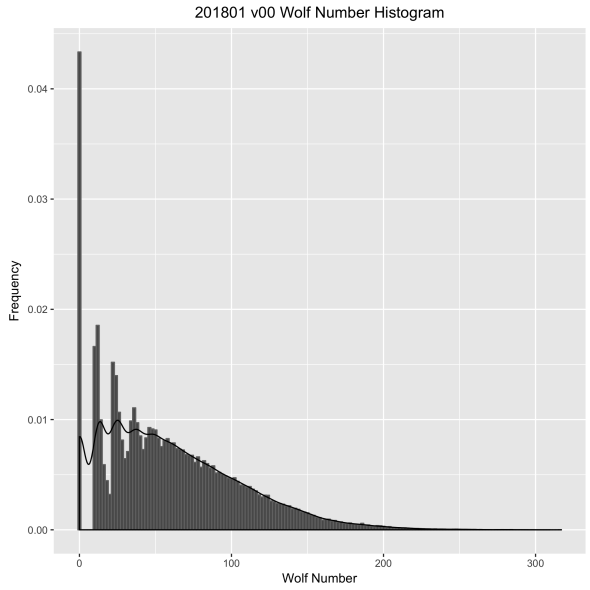

Figure 3: Box plots of raw Wolf number (w) by observer rank.

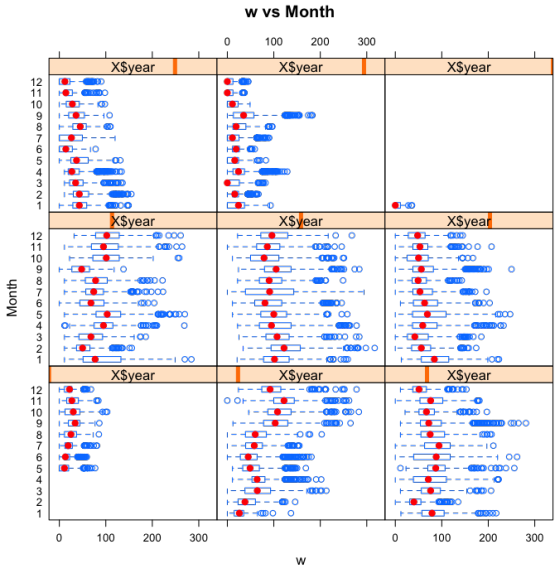

Figure 4: Box plots of raw Wolf number (w) by month and year.

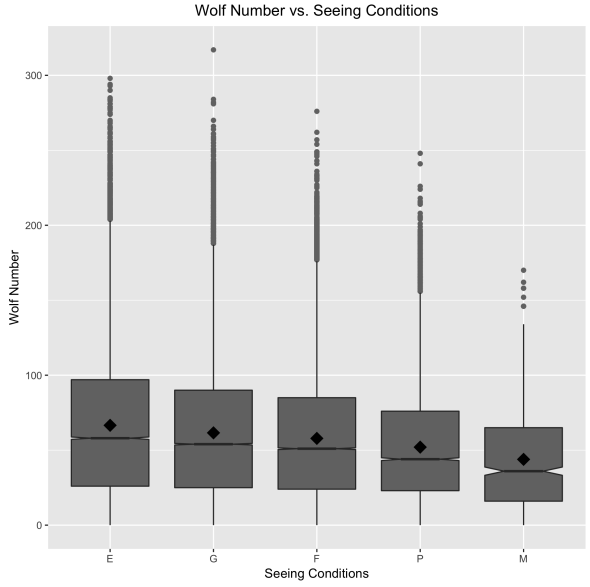

Figure 5: Box plots of raw Wolf number (w) by seeing condition.

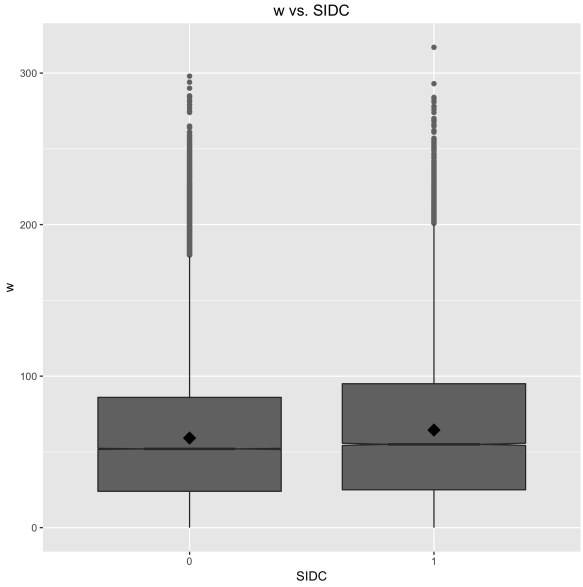

Figure 6: Box plots of raw Wolf number (w) by organization.

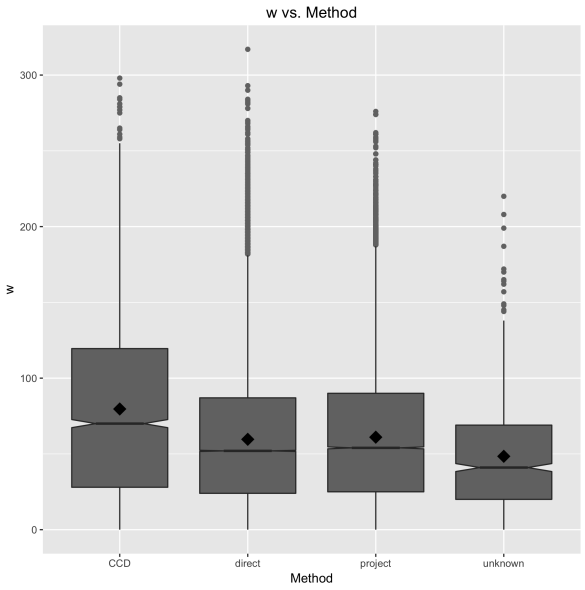

Figure 7: Box plots of raw Wolf number (w) by observer rank.

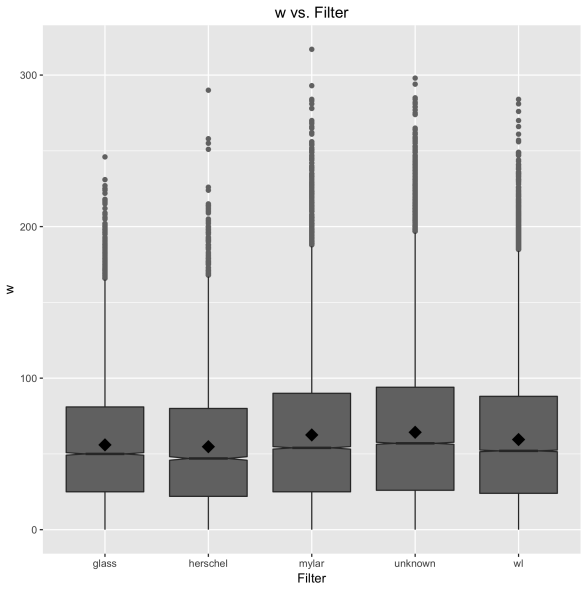

Figure 8: Box plots of raw Wolf number (w) by month and year.

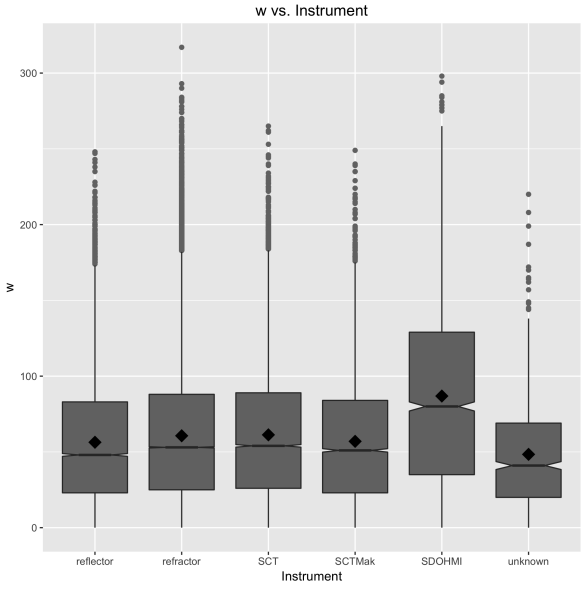

Figure 9: Box plots of raw Wolf number (w) by seeing condition.

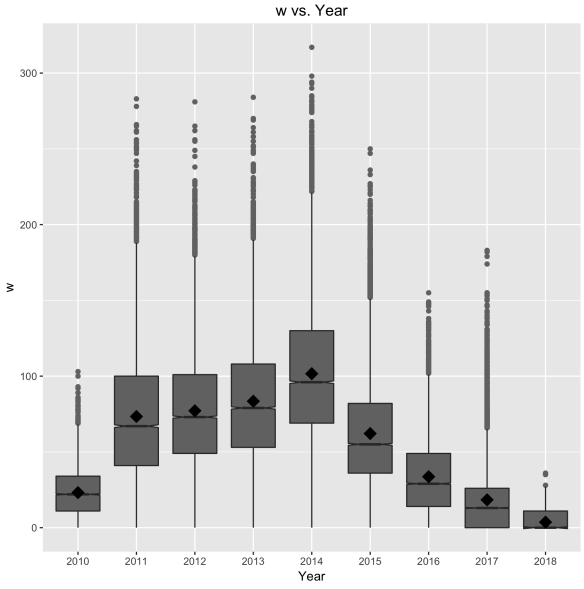

Figure 10: Box plots of raw Wolf number (w) by organization.

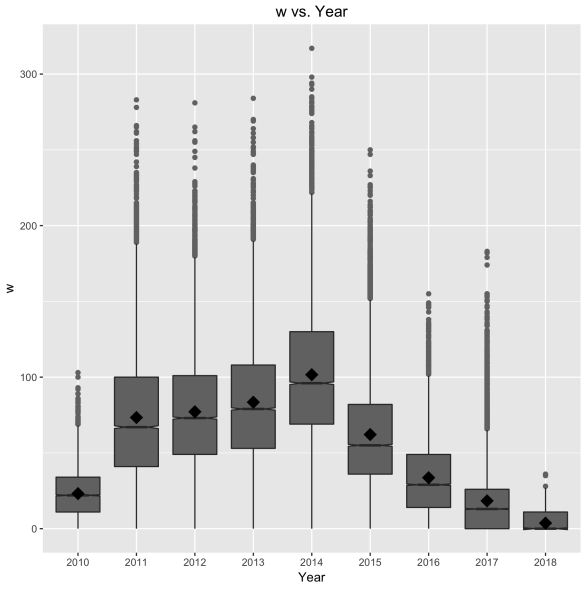

Figure 11: Box plots of raw Wolf number (w) by year.

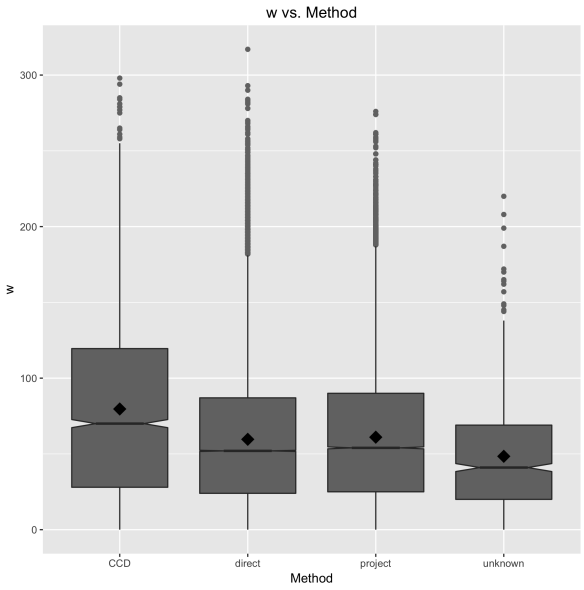

Figure 12: Box plots of raw Wolf number (w) by observing method.