Monthly Report (00)

201807 Data Set

Sunday 12th August, 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.

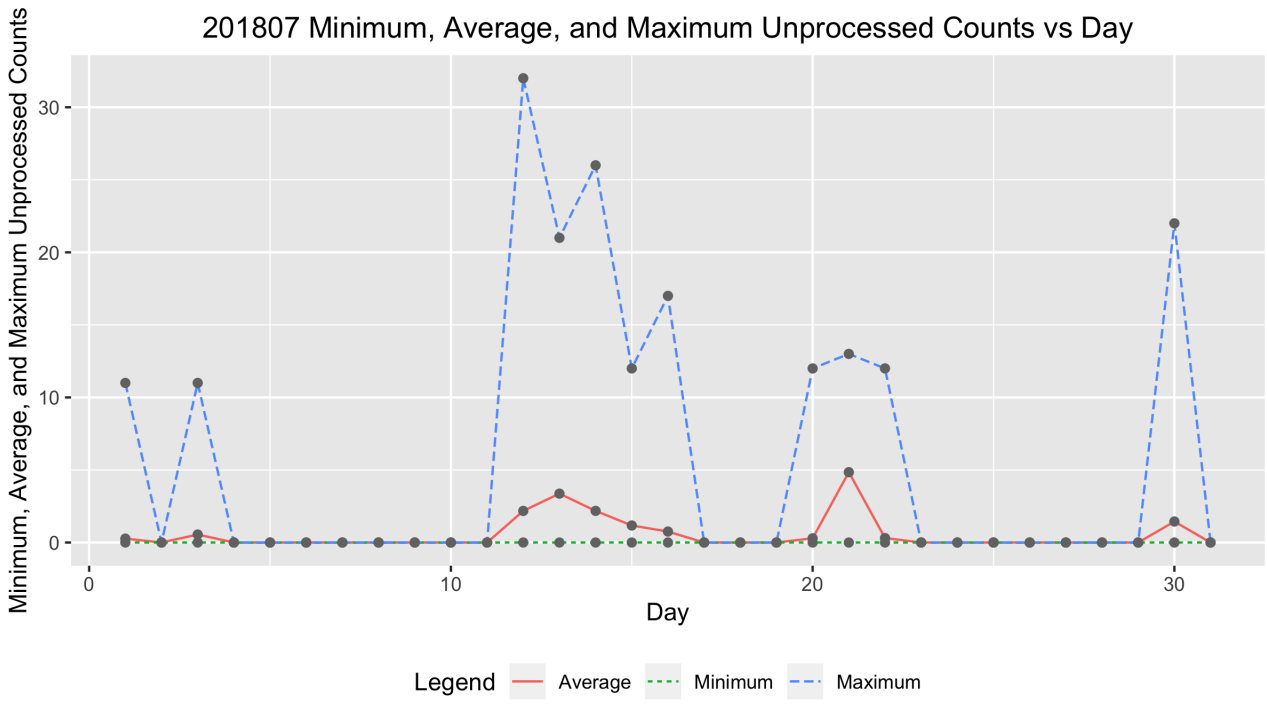

Figure 1: Raw average sunspot count by day of the month.

Table 1: 201807 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	41.0000	0.0000	0.2683	11.0000
2.0000	42.0000	0.0000	0.0000	0.0000
3.0000	40.0000	0.0000	0.5500	11.0000
4.0000	39.0000	0.0000	0.0000	0.0000
5.0000	38.0000	0.0000	0.0000	0.0000
6.0000	36.0000	0.0000	0.0000	0.0000
7.0000	41.0000	0.0000	0.0000	0.0000
8.0000	49.0000	0.0000	0.0000	0.0000
9.0000	43.0000	0.0000	0.0000	0.0000
10.0000	46.0000	0.0000	0.0000	0.0000
11.0000	47.0000	0.0000	0.0000	0.0000
12.0000	45.0000	0.0000	2.1778	32.0000
13.0000	46.0000	0.0000	3.3696	21.0000
14.0000	44.0000	0.0000	2.1818	26.0000
15.0000	40.0000	0.0000	1.1750	12.0000
16.0000	37.0000	0.0000	0.7568	17.0000
17.0000	43.0000	0.0000	0.0000	0.0000
18.0000	44.0000	0.0000	0.0000	0.0000
19.0000	42.0000	0.0000	0.0000	0.0000
20.0000	41.0000	0.0000	0.2927	12.0000
21.0000	33.0000	0.0000	4.8485	13.0000
22.0000	39.0000	0.0000	0.3077	12.0000
23.0000	39.0000	0.0000	0.0000	0.0000
24.0000	33.0000	0.0000	0.0000	0.0000
25.0000	35.0000	0.0000	0.0000	0.0000
26.0000	38.0000	0.0000	0.0000	0.0000
27.0000	42.0000	0.0000	0.0000	0.0000
28.0000	43.0000	0.0000	0.0000	0.0000
29.0000	47.0000	0.0000	0.0000	0.0000
30.0000	38.0000	0.0000	1.4474	22.0000
31.0000	39.0000	0.0000	0.0000	0.0000

3 Error Tables

Data are for the month of July 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
2008.12	2.7705	2.4084	3.1326	0.5000	1.0000
2009.01	5.7966	5.1702	6.4231	1.3000	1.3000
2009.02	5.1095	4.5423	5.6767	0.7000	1.2000
2009.03	6.5681	6.3063	6.8299	0.3000	0.6000
2009.04	7.3876	7.1156	7.6595	0.4000	1.2000
2009.05	7.4922	7.1868	7.7976	1.6000	2.9000
2009.06	6.6605	6.3166	7.0044	3.2000	6.3000
2009.07	6.3437	6.0832	6.6043	3.6000	5.5000
2009.08	7.0073	6.7216	7.2930	0.0000	0.0000
2009.09	7.5799	7.3037	7.8562	4.5000	7.1000
2009.10	7.0696	6.6895	7.4497	4.5000	7.7000
2009.11	7.0246	6.8304	7.2187	3.3000	6.9000
2009.12	6.5385	6.3514	6.7255	10.4000	16.3000
2010.01	21.4069	18.9742	23.8396	13.3000	19.5000
2010.02	17.1505	14.8229	19.4782	19.4000	28.5000

Continued on next page

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

ym	Ra	lci99	uci99	aavso	sidc
2010.03	18.3641	16.0882	20.6399	15.4000	24.0000
2010.04	20.3961	17.9943	22.7979	7.0000	10.4000
2010.05	24.4574	24.0235	24.8912	8.4000	8.7000
2010.06	20.3773	20.0381	20.7165	11.0000	13.6000
2010.07	21.2879	20.9790	21.5968	15.2000	16.1000
2010.08	22.7249	22.3515	23.0982	18.3000	19.6000
2010.09	25.5230	25.1019	25.9441	22.8000	25.2000
2010.10	24.0808	23.6646	24.4970	21.0000	23.5000
2010.11	24.5039	24.0587	24.9491	20.9000	21.6000
2010.12	21.8655	21.4247	22.3064	13.9000	14.5000
2011.01	76.2428	74.6633	77.8224	17.7000	18.7000
2011.02	66.3542	64.9333	67.7751	29.1000	29.6000
2011.03	69.1593	67.8394	70.4793	48.0000	55.8000
2011.04	77.6433	76.2485	79.0382	47.3000	54.4000
2011.05	78.9103	77.5766	80.2441	37.3000	41.5000
2011.06	65.7549	64.6022	66.9076	35.2000	37.0000
2011.07	67.9694	66.8052	69.1336	41.5000	43.8000
2011.08	73.4264	72.2432	74.6097	42.4000	50.5000
2011.09	81.0647	79.6632	82.4662	73.8000	78.0000
2011.10	76.5396	75.2515	77.8277	78.9000	88.0000
2011.11	77.6229	76.0008	79.2451	84.6000	96.7000
2011.12	68.2063	66.8013	69.6113	65.8000	73.0000
2012.01	81.8640	80.2566	83.4714	55.8000	58.2000
2012.02	70.0512	68.6280	71.4744	29.2000	33.1000
2012.03	73.6626	72.3588	74.9664	53.1000	64.1000
2012.04	81.4853	80.0619	82.9087	51.4000	55.2000
2012.05	84.3679	82.9774	85.7584	61.8000	69.0000
2012.06	69.6187	68.4355	70.8019	59.7000	64.5000
2012.07	72.3497	71.1630	73.5364	64.2000	51.3000
2012.08	75.3746	74.1648	76.5844	57.7000	63.1000
2012.09	83.7625	82.3095	85.2156	57.7000	61.5000
2012.10	79.9329	78.4709	81.3948	48.3000	53.3000
2012.11	81.0638	79.4384	82.6893	56.7000	61.4000
2012.12	71.3848	69.8366	72.9331	37.4000	40.8000
2013.01	91.0228	89.2850	92.7606	63.8000	62.9000
2013.02	78.0379	76.4689	79.6070	37.8000	38.0000
2013.03	79.4670	77.8554	81.0786	50.6000	57.9000
2013.04	88.8744	87.3156	90.4333	70.6000	72.4000
2013.05	89.9478	88.3498	91.5457	77.4000	78.7000

Continued on next page

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

ym	Ra	lci99	uci99	aavso	sidc
2013.06	75.7227	74.3855	77.0600	51.0000	52.5000
2013.07	77.7141	76.4678	78.9603	57.0000	57.0000
2013.08	82.5185	81.1936	83.8433	60.0000	66.0000
2013.09	90.4070	88.7972	92.0169	34.6000	36.9000
2013.10	85.2102	83.6333	86.7870	74.5000	85.6000
2013.11	84.8447	82.9460	86.7435	73.9000	77.6000
2013.12	76.8435	75.2064	78.4806	77.8000	90.3000
2014.01	106.0411	103.8142	108.2680	77.4000	82.0000
2014.02	92.7330	90.9096	94.5563	93.9000	102.8000
2014.03	96.6543	94.8980	98.4107	80.9000	92.2000
2014.04	108.2545	106.3682	110.1407	76.9000	84.7000
2014.05	110.2417	108.3639	112.1194	72.3000	75.2000
2014.06	92.5951	91.0311	94.1592	67.2000	71.0000
2014.07	94.7438	93.1693	96.3183	72.5000	72.5000
2014.08	100.7477	99.1755	102.3199	71.2000	74.7000
2014.09	111.5267	109.5634	113.4900	83.2000	87.6000
2014.10	104.8061	102.8788	106.7335	59.5000	60.6000
2014.11	105.3906	103.1994	107.5818	65.8000	71.1000
2014.12	93.4902	91.3265	95.6539	75.8000	78.0000
2015.01	65.5943	64.2831	66.9055	65.9000	67.0000
2015.02	55.9409	54.7095	57.1722	42.4000	44.8000
2015.03	59.0637	57.9827	60.1448	38.0000	38.4000
2015.04	65.7226	64.5426	66.9026	49.0000	54.4000
2015.05	67.2290	66.1151	68.3430	56.3000	58.8000
2015.06	56.3753	55.3566	57.3941	50.2000	68.3000
2015.07	57.1715	56.1937	58.1493	47.9000	65.8000
2015.08	62.0446	61.0001	63.0891	39.5000	57.2000
2015.09	67.7856	66.5520	69.0192	49.2000	72.1000
2015.10	64.2037	62.9626	65.4448	39.3000	48.3000
2015.11	65.2852	63.8672	66.7033	39.6000	55.9000
2015.12	57.9848	56.7105	59.2592	36.4000	44.8000
2016.01	35.8949	35.1457	36.6441	33.7000	43.3000
2016.02	30.7607	30.1195	31.4018	38.3000	46.8000
2016.03	31.9585	31.3208	32.5961	30.5000	38.9000
2016.04	35.4523	34.7758	36.1288	26.6000	30.9000
2016.05	36.3575	35.6917	37.0233	33.7000	48.4000
2016.06	30.1574	29.6422	30.6726	13.1000	19.5000
2016.07	31.1532	30.6511	31.6552	21.2000	27.5000
2016.08	33.4362	32.8468	34.0256	33.0000	47.9000

Continued on next page

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

ym	Ra	lci99	uci99	aavso	sidc
2016.09	37.3953	36.7112	38.0793	27.7000	37.1000
2016.10	35.0177	34.3412	35.6942	22.7000	31.7000
2016.11	35.2023	34.4656	35.9389	14.0000	22.2000
2016.12	31.6706	30.9942	32.3469	11.1000	20.0000
2017.01	19.5301	19.1193	19.9410	18.4000	26.2000
2017.02	16.8006	16.4318	17.1695	14.4000	20.6000
2017.03	17.6090	17.2736	17.9445	11.3000	15.5000
2017.04	19.7518	19.4029	20.1006	21.6000	33.2000
2017.05	19.9262	19.5816	20.2708	12.5000	18.1000
2017.06	16.5444	16.2680	16.8209	15.5000	19.3000
2017.07	17.1654	16.8922	17.4387	11.5000	16.3000
2017.08	18.3549	18.0393	18.6706	22.8000	35.7000
2017.09	20.7869	20.3837	21.1902	34.6000	42.9000
2017.10	19.0513	18.6730	19.4296	10.5000	11.0000
2017.11	19.0009	18.6032	19.3986	4.2000	5.6000
2017.12	16.9706	16.7106	17.2306	4.0000	4.6000
2018.01	6.2351	6.1017	6.3685	3.1000	6.3000
2018.02	5.3122	5.1849	5.4394	6.8000	11.8000
2018.03	5.4706	5.3607	5.5805	1.1000	1.2000
2018.04	6.0653	5.9439	6.1867	4.7000	7.5000
2018.05	6.2605	6.1401	6.3810	8.4000	14.0000
2018.06	5.1727	5.0769	5.2684	10.2000	13.6000
2018.07	5.3377	5.2744	5.4010	0.5000	1.7000

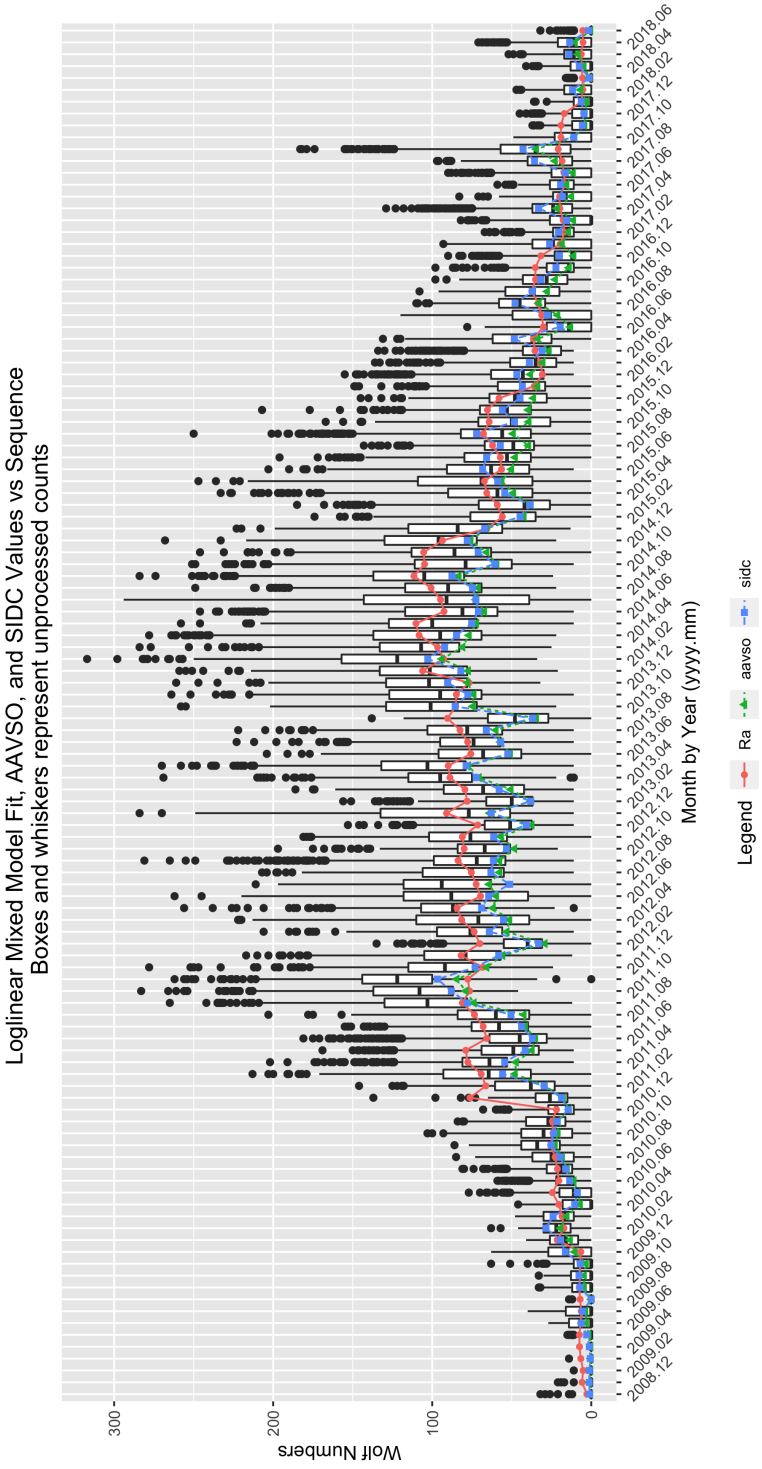

Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201807 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	1.4205	0.3185	4.4599	0.0000
seeF	-0.2185	0.0060	-36.1877	0.0000
seeG	-0.1154	0.0053	-21.9508	0.0000
seeM	-0.1951	0.0247	-7.8870	0.0000
seeP	-0.3251	0.0086	-37.5802	0.0000
sidc1	0.1506	0.0698	2.1575	0.0310
year2009	0.6441	0.3195	2.0161	0.0438
year2010	1.8548	0.3173	5.8461	0.0000
year2011	2.9758	0.3172	9.3822	0.0000
year2012	3.0134	0.3172	9.5010	0.0000
year2013	3.1094	0.3172	9.8039	0.0000
year2014	3.3065	0.3172	10.4252	0.0000
year2015	2.8219	0.3172	8.8972	0.0000
year2016	2.2049	0.3172	6.9511	0.0000
year2017	1.6002	0.3172	5.0441	0.0000
year2018	0.4491	0.3177	1.4134	0.1575
mon2	-0.1450	0.0095	-15.2380	0.0000
mon3	-0.1113	0.0089	-12.4462	0.0000
mon4	-0.0088	0.0086	-1.0214	0.3070
mon5	0.0035	0.0085	0.4116	0.6806
mon6	-0.1810	0.0089	-20.4384	0.0000
mon7	-0.1565	0.0086	-18.1763	0.0000
mon8	-0.0867	0.0085	-10.2533	0.0000
mon9	0.0230	0.0085	2.7126	0.0067
mon10	-0.0330	0.0087	-3.7903	0.0002
mon11	-0.0127	0.0091	-1.3921	0.1639
mon12	-0.1225	0.0093	-13.2014	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: 201807 Summary of Sunspot Numbers

year	mon	day	obs	sidc
Min. :2008	Min. : 1.00	Min. : 0.00	Length:102738	Min. :0.0000
1st Qu.:2012	1st Qu.: 4.00	1st Qu.: 8.00	Class :character	1st Qu.:0.0000
Median :2014	Median : 7.00	Median :16.00	Mode :character	Median :0.0000
Mean :2014	Mean : 6.57	Mean :15.73		Mean :0.2632
3rd Qu.:2016	3rd Qu.: 9.00	3rd Qu.:23.00		3rd Qu.:1.0000
Max. :2018	Max. :12.00	Max. :31.00		Max. :1.0000

Table 5: 201807 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.00	Min. : 0.00	Min. : 0.00	Length:102738	Length:102738
1st Qu.: 1.00	1st Qu.: 4.00	1st Qu.: 17.00	Class :character	Class :character
Median : 3.00	Median : 14.00	Median : 47.00	Mode :character	Mode :character
Mean : 3.47	Mean : 20.79	Mean : 55.49		
3rd Qu.: 5.00	3rd Qu.: 31.00	3rd Qu.: 84.00		
Max. :19.00	Max. :204.00	Max. :317.00		

Table 6: 201807 Summary of Sunspot Numbers

inst	filter	unit
Length:102738	Length:102738	Length:102738
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201807 Summary of Sunspot Numbers

aperture	eyep	foclen	mag
Min. : 0.0	Min. : 0.00	Min. : 0	Min. : 0.0
1st Qu.: 76.0	1st Qu.: 2.00	1st Qu.: 800	1st Qu.: 40.0
Median : 90.0	Median : 13.00	Median :1000	Median : 57.5
Mean : 110.2	Mean : 19.85	Mean :1107	Mean : 186.7
3rd Qu.: 125.0	3rd Qu.: 23.00	3rd Qu.:1283	3rd Qu.: 76.0
Max. :1524.0	Max. :2010.00	Max. :4300	Max. :4591.0

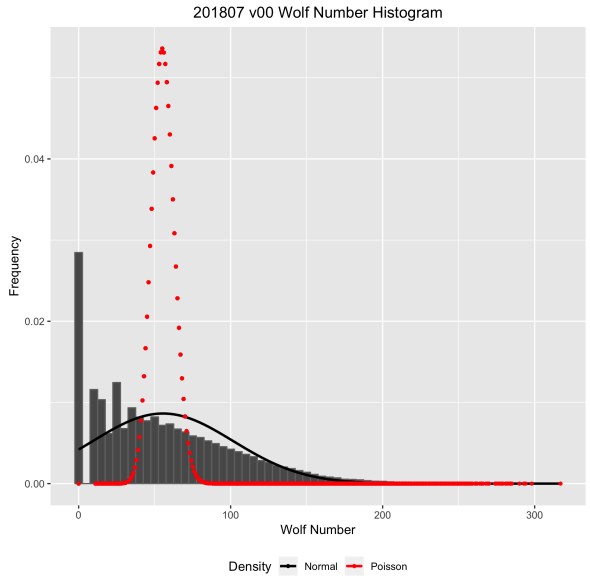

Figure 3: Box plots of raw Wolf number (w) by observer rank.

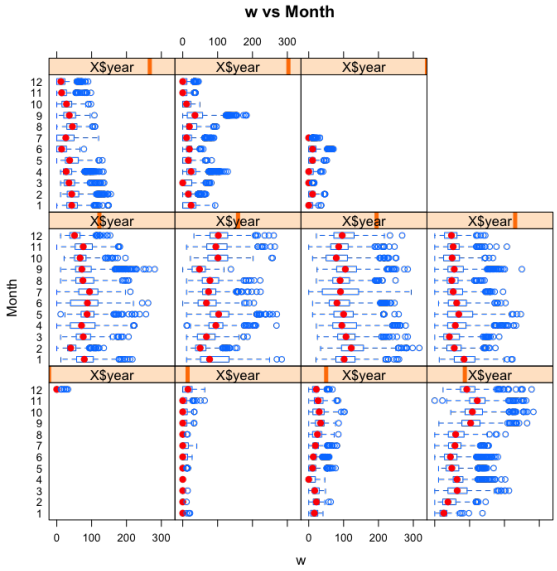

Figure 4: Box plots of raw Wolf number (w) by month and year.

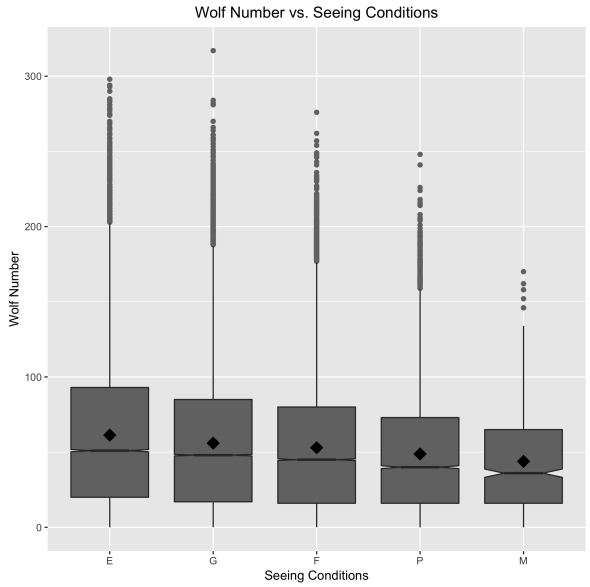

Figure 5: Box plots of raw Wolf number (w) by seeing condition.

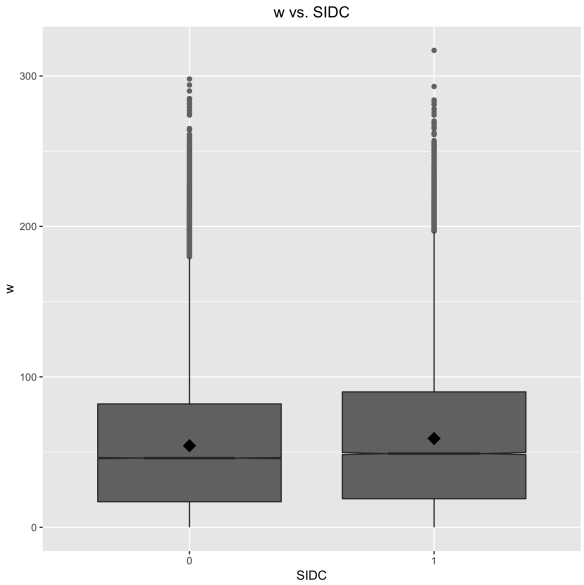

Figure 6: Box plots of raw Wolf number (w) by organization.

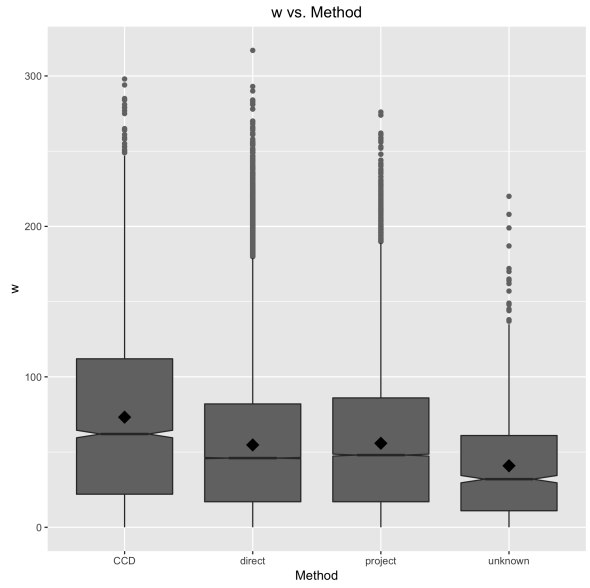

Figure 7: Box plots of raw Wolf number (w) by observer rank.

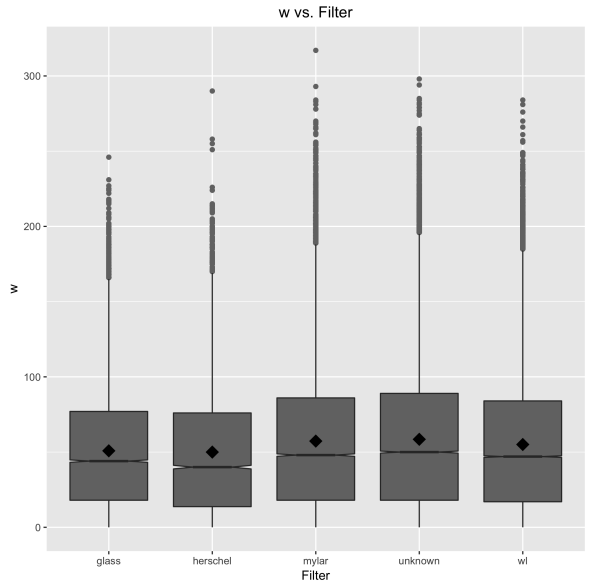

Figure 8: Box plots of raw Wolf number (w) by month and year.

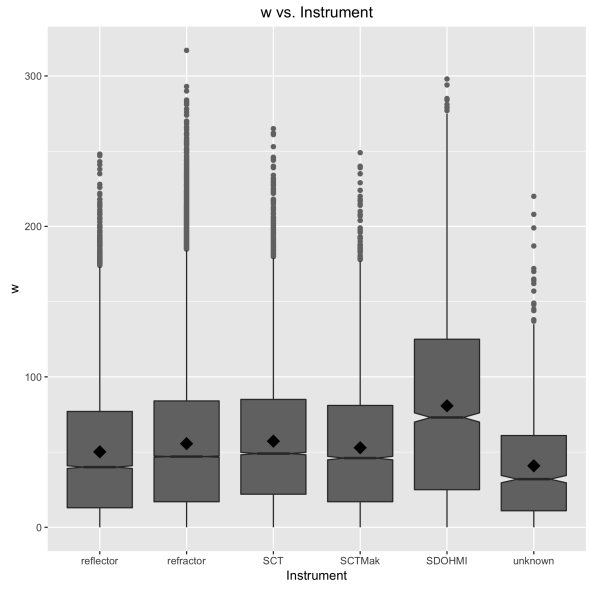

Figure 9: Box plots of raw Wolf number (w) by seeing condition.

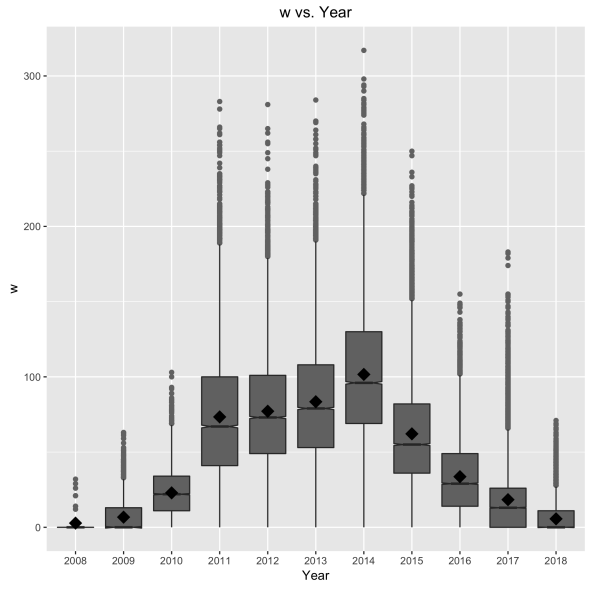

Figure 10: Box plots of raw Wolf number (w) by organization.

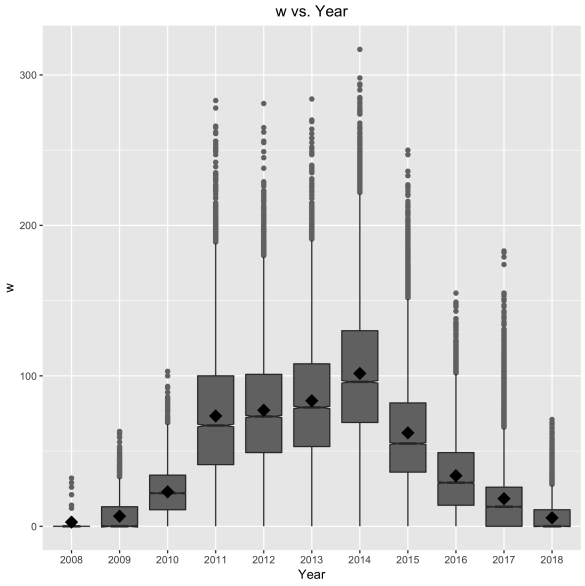

Figure 11: Box plots of raw Wolf number (w) by year.

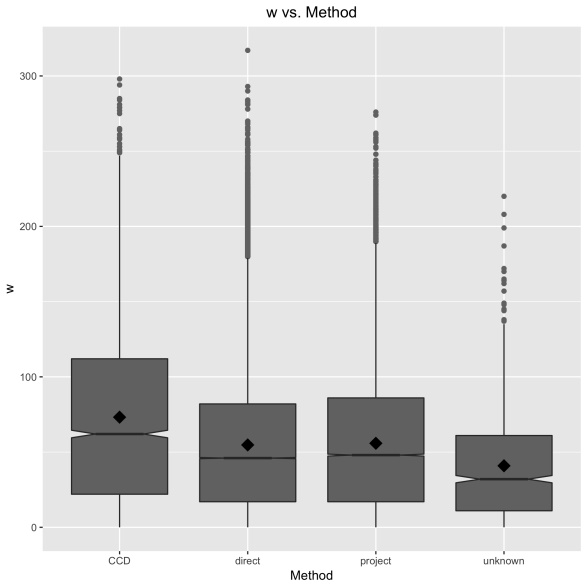

Figure 12: Box plots of raw Wolf number (w) by observing method.