

Monthly Report (00)

201809 Data Set

Tuesday 16th October, 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.

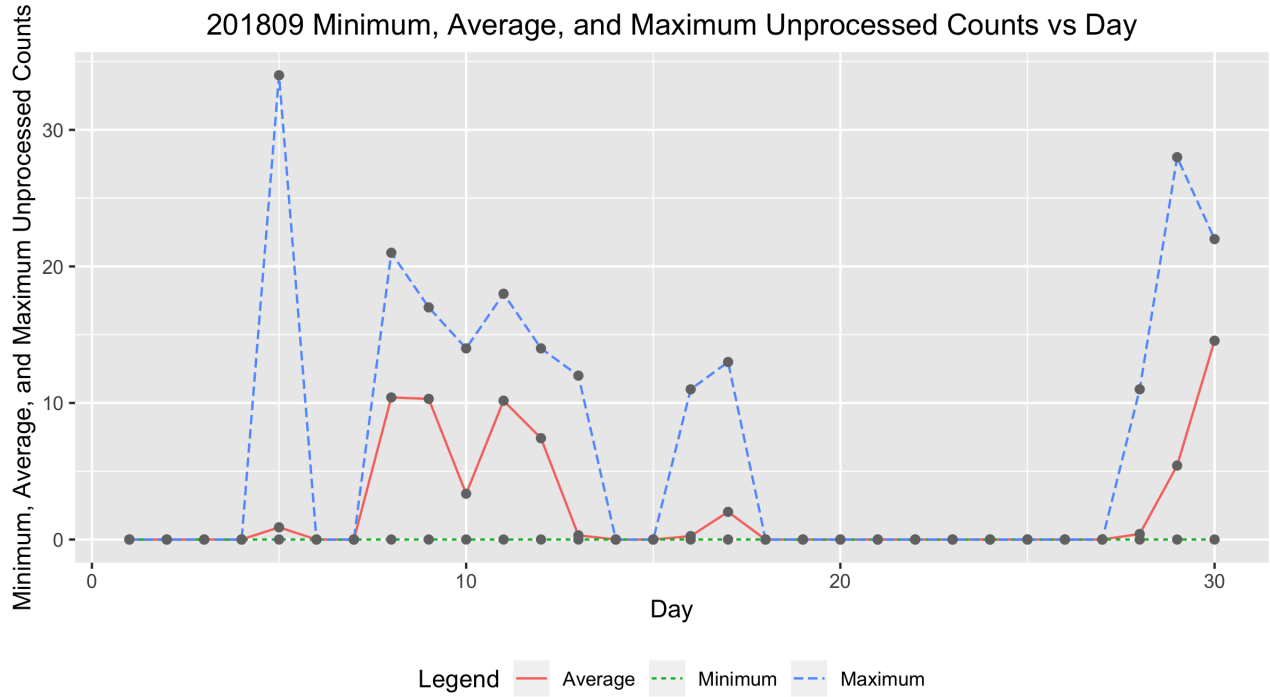


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201809 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	36.0000	0.0000	0.0000	0.0000
2.0000	41.0000	0.0000	0.0000	0.0000
3.0000	41.0000	0.0000	0.0000	0.0000
4.0000	36.0000	0.0000	0.0000	0.0000
5.0000	38.0000	0.0000	0.8947	34.0000
6.0000	38.0000	0.0000	0.0000	0.0000
7.0000	35.0000	0.0000	0.0000	0.0000
8.0000	37.0000	0.0000	10.4054	21.0000
9.0000	40.0000	0.0000	10.3000	17.0000
10.0000	34.0000	0.0000	3.3529	14.0000
11.0000	30.0000	0.0000	10.1667	18.0000
12.0000	31.0000	0.0000	7.4194	14.0000
13.0000	39.0000	0.0000	0.3077	12.0000
14.0000	35.0000	0.0000	0.0000	0.0000
15.0000	37.0000	0.0000	0.0000	0.0000
16.0000	44.0000	0.0000	0.2500	11.0000
17.0000	38.0000	0.0000	2.0263	13.0000
18.0000	34.0000	0.0000	0.0000	0.0000
19.0000	40.0000	0.0000	0.0000	0.0000
20.0000	31.0000	0.0000	0.0000	0.0000
21.0000	32.0000	0.0000	0.0000	0.0000
22.0000	36.0000	0.0000	0.0000	0.0000
23.0000	41.0000	0.0000	0.0000	0.0000
24.0000	39.0000	0.0000	0.0000	0.0000
25.0000	30.0000	0.0000	0.0000	0.0000
26.0000	32.0000	0.0000	0.0000	0.0000
27.0000	33.0000	0.0000	0.0000	0.0000
28.0000	27.0000	0.0000	0.4074	11.0000
29.0000	36.0000	0.0000	5.4167	28.0000
30.0000	34.0000	0.0000	14.5588	22.0000

3 Error Tables

Data are for the month of September 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 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
2008.12	2.7705	2.4080	3.1330	0.5000	1.0000
2009.01	5.7991	5.1716	6.4267	1.3000	1.3000
2009.02	5.1121	4.5438	5.6803	0.7000	1.2000
2009.03	6.5728	6.3112	6.8344	0.3000	0.6000
2009.04	7.3925	7.1208	7.6641	0.4000	1.2000
2009.05	7.4975	7.1925	7.8026	1.6000	2.9000
2009.06	6.6645	6.3207	7.0082	3.2000	6.3000
2009.07	6.3503	6.0898	6.6108	3.6000	5.5000
2009.08	7.0390	6.7527	7.3254	0.0000	0.0000
2009.09	7.5262	7.2522	7.8001	4.5000	7.1000
2009.10	7.0708	6.6910	7.4506	4.5000	7.7000
2009.11	7.0242	6.8300	7.2184	3.3000	6.9000
2009.12	6.5374	6.3504	6.7244	10.4000	16.3000
2010.01	21.4264	18.9884	23.8644	13.3000	19.5000
2010.02	17.1633	14.8307	19.4959	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	18.3783	16.0979	20.6587	15.4000	24.0000
2010.04	20.4111	18.0047	22.8176	7.0000	10.4000
2010.05	24.4737	24.0394	24.9079	8.4000	8.7000
2010.06	20.3914	20.0520	20.7308	11.0000	13.6000
2010.07	21.3087	20.9994	21.6179	15.2000	16.1000
2010.08	22.8270	22.4520	23.2020	18.3000	19.6000
2010.09	25.3422	24.9241	25.7603	22.8000	25.2000
2010.10	24.0868	23.6705	24.5031	21.0000	23.5000
2010.11	24.5107	24.0654	24.9560	20.9000	21.6000
2010.12	21.8705	21.4294	22.3116	13.9000	14.5000
2011.01	76.2880	74.7084	77.8677	17.7000	18.7000
2011.02	66.3982	64.9784	67.8180	29.1000	29.6000
2011.03	69.1922	67.8713	70.5131	48.0000	55.8000
2011.04	77.6812	76.2853	79.0770	47.3000	54.4000
2011.05	78.9463	77.6113	80.2813	37.3000	41.5000
2011.06	65.7825	64.6288	66.9362	35.2000	37.0000
2011.07	68.0201	66.8547	69.1855	41.5000	43.8000
2011.08	73.7405	72.5517	74.9294	42.4000	50.5000
2011.09	80.4877	79.0947	81.8807	73.8000	78.0000
2011.10	76.5414	75.2526	77.8301	78.9000	88.0000
2011.11	77.6263	76.0036	79.2490	84.6000	96.7000
2011.12	68.2030	66.7975	69.6085	65.8000	73.0000
2012.01	81.9043	80.2955	83.5131	55.8000	58.2000
2012.02	70.0841	68.6599	71.5084	29.2000	33.1000
2012.03	73.6954	72.3905	75.0002	53.1000	64.1000
2012.04	81.5199	80.0955	82.9443	51.4000	55.2000
2012.05	84.4094	83.0179	85.8010	61.8000	69.0000
2012.06	69.6488	68.4650	70.8327	59.7000	64.5000
2012.07	72.4101	71.2222	73.5980	64.2000	51.3000
2012.08	75.6991	74.4839	76.9143	57.7000	63.1000
2012.09	83.1540	81.7112	84.5969	57.7000	61.5000
2012.10	79.9329	78.4707	81.3951	48.3000	53.3000
2012.11	81.0638	79.4380	82.6896	56.7000	61.4000
2012.12	71.3830	69.8343	72.9316	37.4000	40.8000
2013.01	91.0633	89.3241	92.8025	63.8000	62.9000
2013.02	78.0771	76.5069	79.6473	37.8000	38.0000
2013.03	79.5049	77.8916	81.1182	50.6000	57.9000
2013.04	88.9102	87.3498	90.4706	70.6000	72.4000
2013.05	89.9856	88.3864	91.5848	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	75.7534	74.4152	77.0917	51.0000	52.5000
2013.07	77.7790	76.5313	79.0267	57.0000	57.0000
2013.08	82.8724	81.5416	84.2033	60.0000	66.0000
2013.09	89.7459	88.1472	91.3446	34.6000	36.9000
2013.10	85.2096	83.6320	86.7873	74.5000	85.6000
2013.11	84.8427	82.9432	86.7422	73.9000	77.6000
2013.12	76.8375	75.1996	78.4755	77.8000	90.3000
2014.01	106.0855	103.8563	108.3147	77.4000	82.0000
2014.02	92.7756	90.9505	94.6007	93.9000	102.8000
2014.03	96.6944	94.9363	98.4525	80.9000	92.2000
2014.04	108.2961	106.4081	110.1841	76.9000	84.7000
2014.05	110.2805	108.4009	112.1602	72.3000	75.2000
2014.06	92.6298	91.0647	94.1949	67.2000	71.0000
2014.07	94.8183	93.2418	96.3948	72.5000	72.5000
2014.08	101.1810	99.6015	102.7606	71.2000	74.7000
2014.09	110.7153	108.7654	112.6652	83.2000	87.6000
2014.10	104.8046	102.8764	106.7327	59.5000	60.6000
2014.11	105.3849	103.1925	107.5773	65.8000	71.1000
2014.12	93.4835	91.3187	95.6482	75.8000	78.0000
2015.01	65.6211	64.3082	66.9340	65.9000	67.0000
2015.02	55.9654	54.7328	57.1980	42.4000	44.8000
2015.03	59.0932	58.0111	60.1753	38.0000	38.4000
2015.04	65.7495	64.5693	66.9296	49.0000	54.4000
2015.05	67.2597	66.1451	68.3743	56.3000	58.8000
2015.06	56.4005	55.3815	57.4196	50.2000	68.3000
2015.07	57.2222	56.2437	58.2008	47.9000	65.8000
2015.08	62.3169	61.2676	63.3662	39.5000	57.2000
2015.09	67.2948	66.0699	68.5196	49.2000	72.1000
2015.10	64.2024	62.9608	65.4440	39.3000	48.3000
2015.11	65.2825	63.8635	66.7014	39.6000	55.9000
2015.12	57.9821	56.7072	59.2570	36.4000	44.8000
2016.01	35.9101	35.1602	36.6600	33.7000	43.3000
2016.02	30.7732	30.1314	31.4150	38.3000	46.8000
2016.03	31.9714	31.3331	32.6097	30.5000	38.9000
2016.04	35.4673	34.7903	36.1443	26.6000	30.9000
2016.05	36.3744	35.7081	37.0408	33.7000	48.4000
2016.06	30.1683	29.6527	30.6840	13.1000	19.5000
2016.07	31.1785	30.6760	31.6810	21.2000	27.5000
2016.08	33.5805	32.9885	34.1726	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	37.1231	36.4439	37.8023	27.7000	37.1000
2016.10	35.0165	34.3397	35.6932	22.7000	31.7000
2016.11	35.2032	34.4663	35.9401	14.0000	22.2000
2016.12	31.6698	30.9932	32.3465	11.1000	20.0000
2017.01	19.5369	19.1258	19.9479	18.4000	26.2000
2017.02	16.8080	16.4388	17.1772	14.4000	20.6000
2017.03	17.6164	17.2806	17.9521	11.3000	15.5000
2017.04	19.7587	19.4097	20.1077	21.6000	33.2000
2017.05	19.9340	19.5893	20.2788	12.5000	18.1000
2017.06	16.5487	16.2721	16.8253	15.5000	19.3000
2017.07	17.1767	16.9032	17.4502	11.5000	16.3000
2017.08	18.4340	18.1164	18.7516	22.8000	35.7000
2017.09	20.6470	20.2420	21.0519	34.6000	42.9000
2017.10	19.0498	18.6700	19.4296	10.5000	11.0000
2017.11	19.0025	18.6051	19.3999	4.2000	5.6000
2017.12	16.9791	16.7196	17.2385	4.0000	4.6000
2018.01	6.0098	5.8797	6.1399	3.1000	6.3000
2018.02	5.1175	4.9928	5.2421	6.8000	11.8000
2018.03	5.2683	5.1612	5.3753	1.1000	1.2000
2018.04	5.8538	5.7347	5.9728	4.7000	7.5000
2018.05	6.0522	5.9321	6.1724	8.4000	14.0000
2018.06	5.0087	4.9147	5.1028	10.2000	13.6000
2018.07	5.1716	5.1102	5.2330	0.5000	1.7000
2018.08	5.4969	5.3945	5.5993	5.9000	9.5000
2018.09	5.9350	5.8183	6.0517	1.6000	2.9000

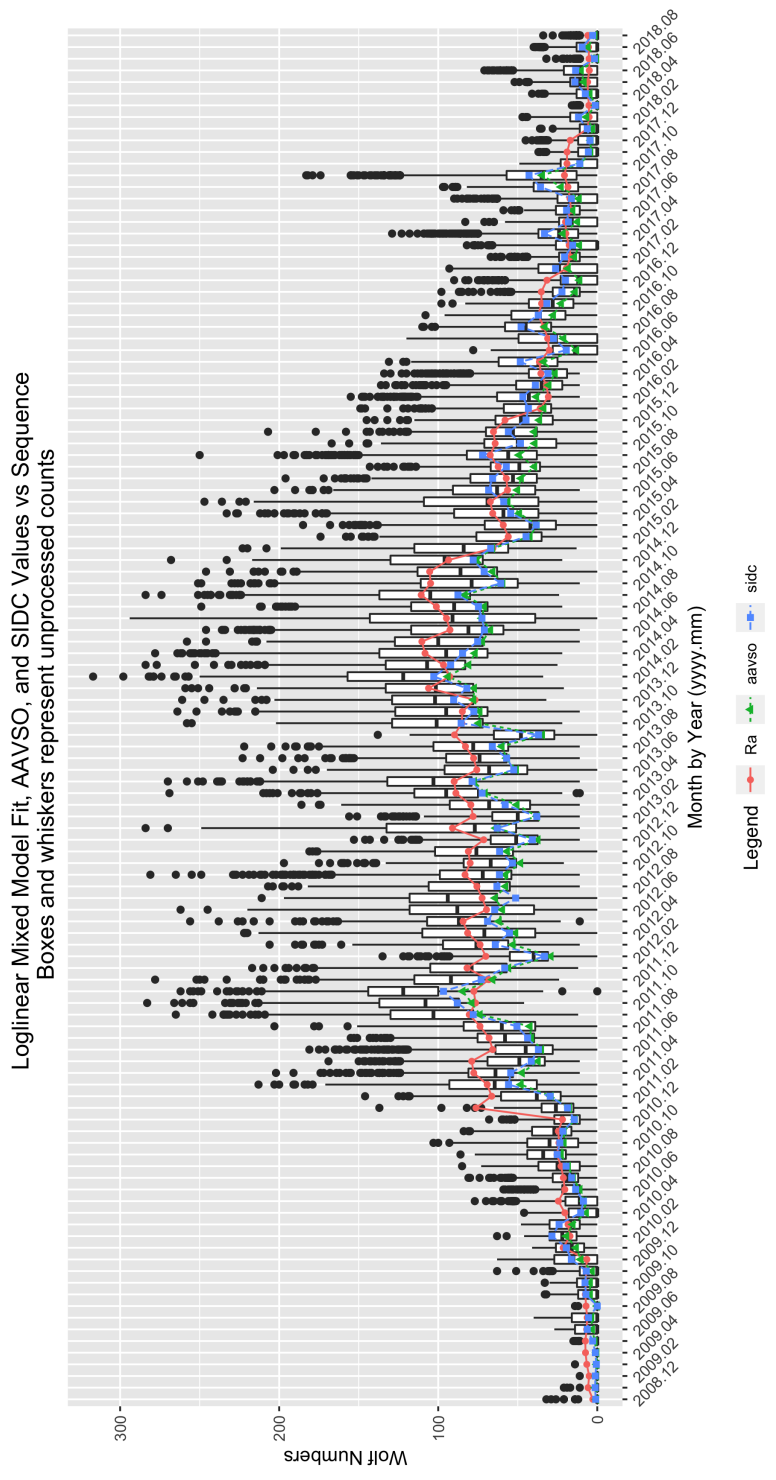


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201809 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4226	0.3176	4.4797	0.0000
seeF	-0.2186	0.0060	-36.3523	0.0000
seeG	-0.1157	0.0052	-22.1003	0.0000
seeM	-0.1962	0.0247	-7.9567	0.0000
seeP	-0.3253	0.0086	-37.7491	0.0000
sidc1	0.1575	0.0701	2.2470	0.0246
year2009	0.6440	0.3185	2.0218	0.0432
year2010	1.8551	0.3163	5.8643	0.0000
year2011	2.9758	0.3162	9.4103	0.0000
year2012	3.0134	0.3162	9.5293	0.0000
year2013	3.1094	0.3162	9.8331	0.0000
year2014	3.3064	0.3162	10.4561	0.0000
year2015	2.8219	0.3162	8.9234	0.0000
year2016	2.2049	0.3163	6.9715	0.0000
year2017	1.6000	0.3163	5.0583	0.0000
year2018	0.4128	0.3167	1.3035	0.1924
mon2	-0.1449	0.0095	-15.2821	0.0000
mon3	-0.1113	0.0089	-12.4830	0.0000
mon4	-0.0089	0.0086	-1.0292	0.3034
mon5	0.0035	0.0084	0.4139	0.6790
mon6	-0.1811	0.0088	-20.5051	0.0000
mon7	-0.1562	0.0086	-18.1945	0.0000
mon8	-0.0829	0.0084	-9.8662	0.0000
mon9	0.0152	0.0084	1.7977	0.0722
mon10	-0.0335	0.0087	-3.8585	0.0001
mon11	-0.0131	0.0091	-1.4475	0.1478
mon12	-0.1230	0.0092	-13.2964	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: 201809 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 0.00	Length:105019	Min. :0.0000
1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00	Class :character	1st Qu.:0.0000
Median :2014	Median : 7.000	Median :16.00	Mode :character	Median :0.0000
Mean :2014	Mean : 6.611	Mean :15.73		Mean :0.2628
3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.00		3rd Qu.:1.0000
Max. :2018	Max. :12.000	Max. :31.00		Max. :1.0000

Table 5: 201809 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:105019	Length:105019
1st Qu.: 1.000	1st Qu.: 3.00	1st Qu.: 15.00	Class :character	Class :character
Median : 3.000	Median : 13.00	Median : 46.00	Mode :character	Mode :character
Mean : 3.403	Mean : 20.36	Mean : 54.39		
3rd Qu.: 5.000	3rd Qu.: 30.00	3rd Qu.: 83.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201809 Summary of Sunspot Numbers

inst	filter	unit
Length:105019	Length:105019	Length:105019
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201809 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.: 700	1st Qu.: 40.0
Median : 90.0	Median : 13.00	Median :1000	Median : 57.5
Mean : 108.1	Mean : 20.93	Mean :1083	Mean : 186.4
3rd Qu.: 125.0	3rd Qu.: 23.00	3rd Qu.:1260	3rd Qu.: 76.0
Max. :1524.0	Max. :2010.00	Max. :4300	Max. :4591.0

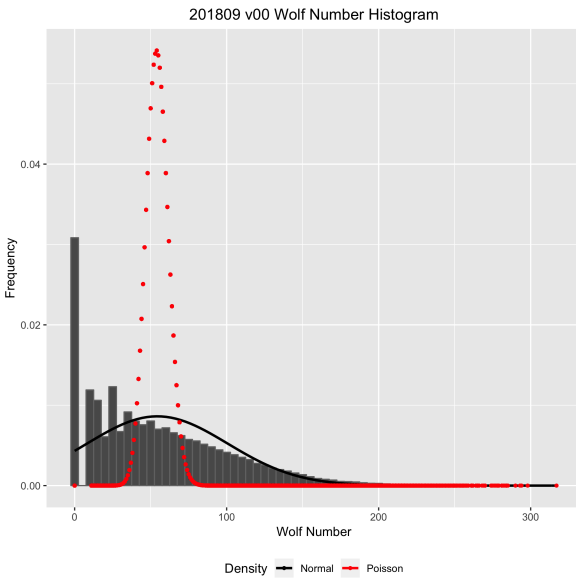


Figure 3: Box plots of raw Wolf number (w) by observer rank.

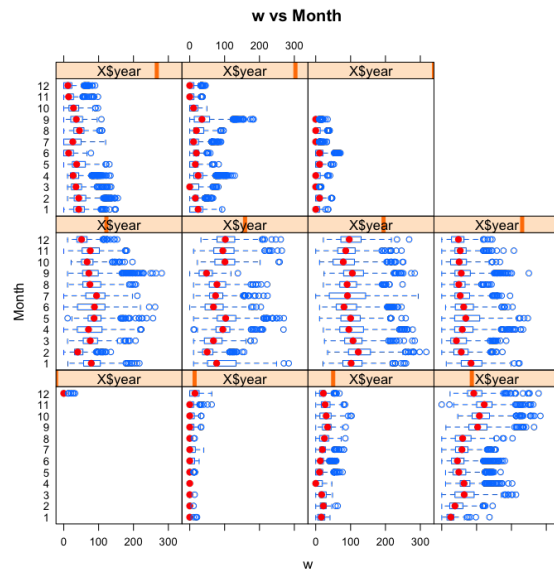


Figure 4: Box plots of raw Wolf number (w) by month and year.

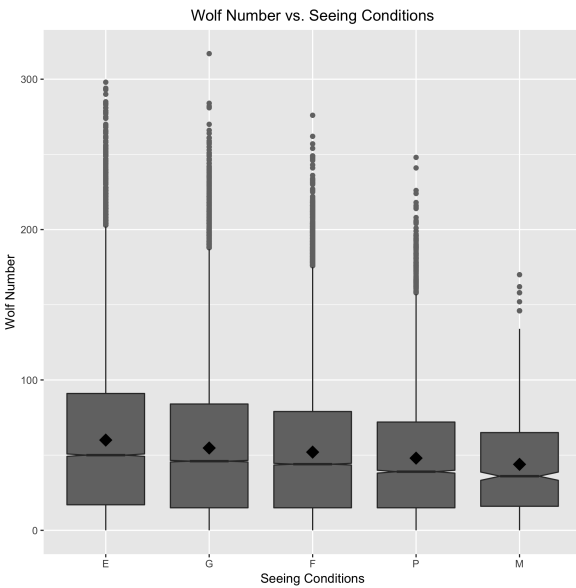


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

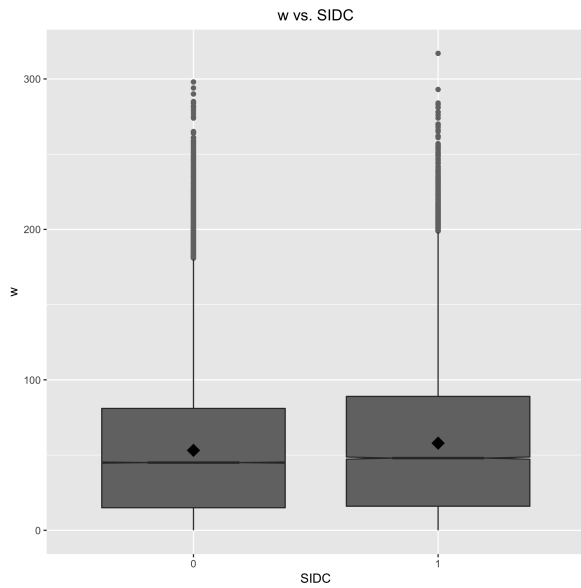


Figure 6: Box plots of raw Wolf number (w) by organization.

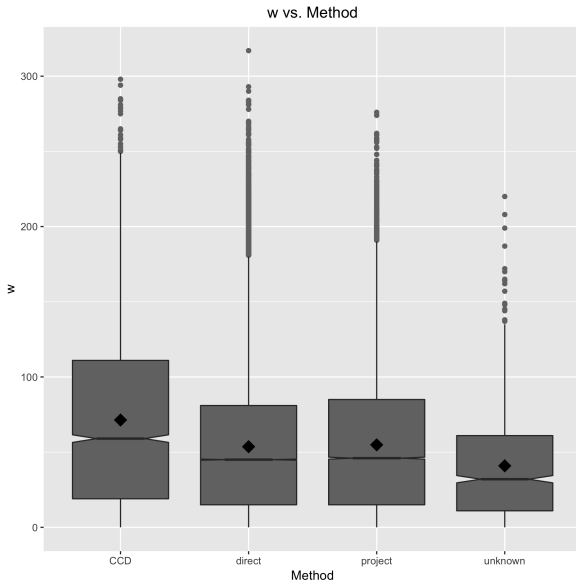


Figure 7: Box plots of raw Wolf number (w) by observer rank.

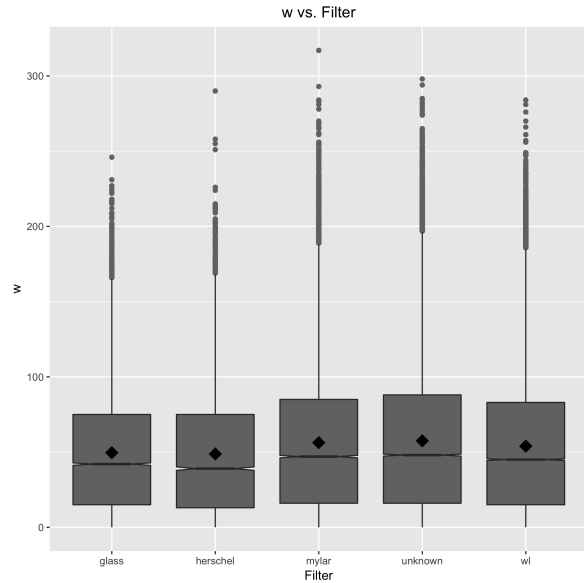


Figure 8: Box plots of raw Wolf number (w) by month and year.

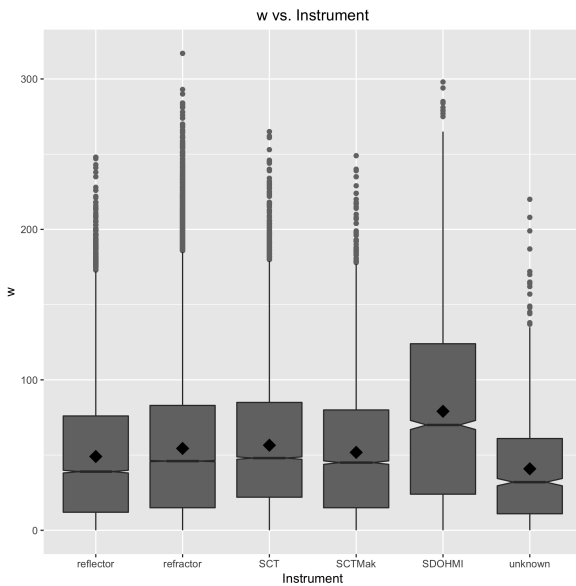


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

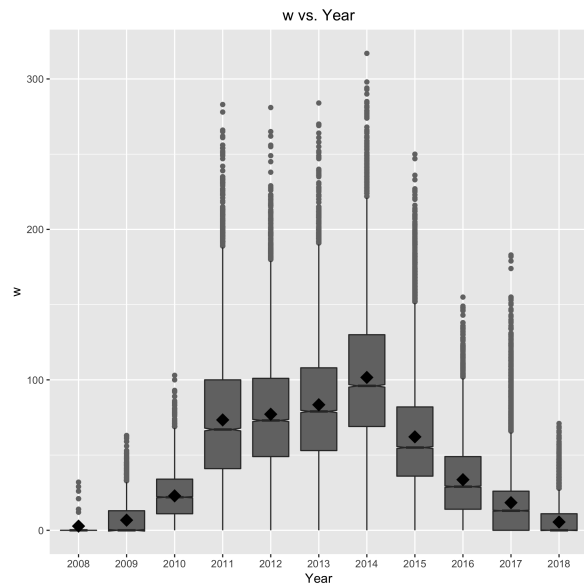


Figure 10: Box plots of raw Wolf number (w) by organization.

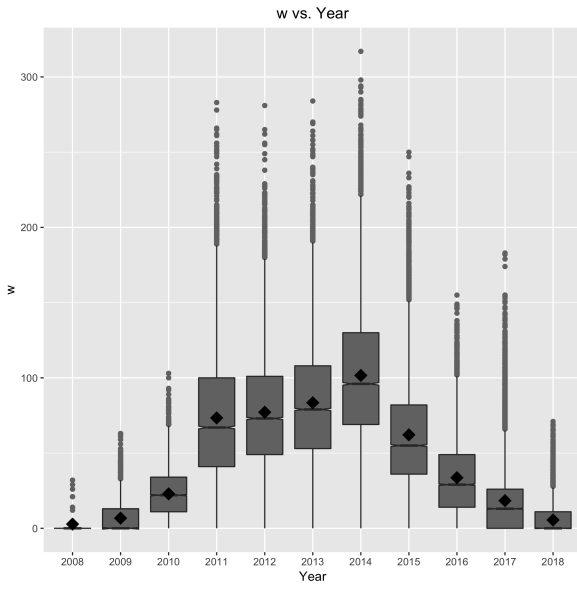


Figure 11: Box plots of raw Wolf number (w) by year.

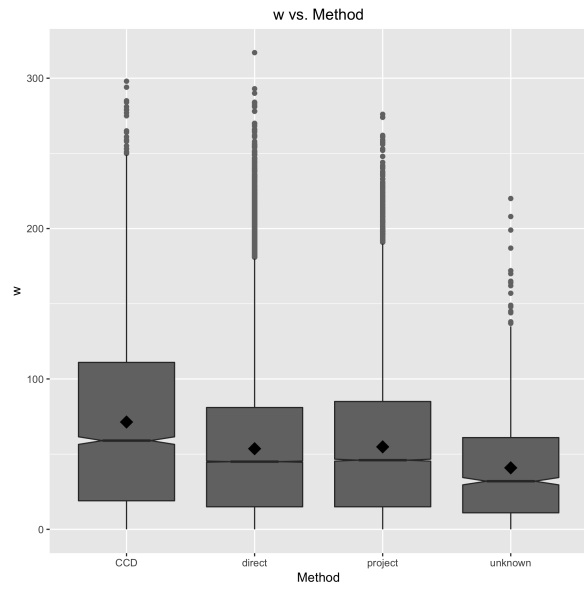


Figure 12: Box plots of raw Wolf number (w) by observing method.