

Monthly Report (00)
201808 Data Set

Wednesday 12th September, 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.

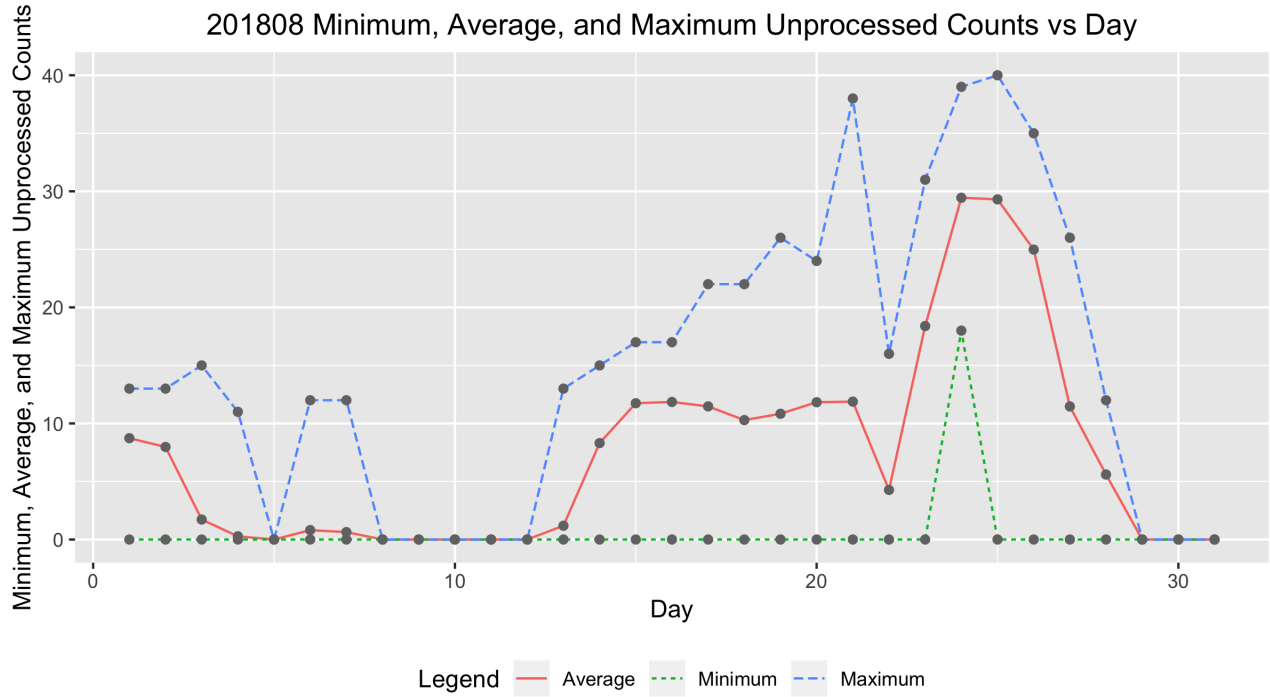


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201808 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	37.0000	0.0000	8.7297	13.0000
2.0000	39.0000	0.0000	7.9744	13.0000
3.0000	36.0000	0.0000	1.7222	15.0000
4.0000	40.0000	0.0000	0.2750	11.0000
5.0000	43.0000	0.0000	0.0000	0.0000
6.0000	42.0000	0.0000	0.8095	12.0000
7.0000	36.0000	0.0000	0.6389	12.0000
8.0000	41.0000	0.0000	0.0000	0.0000
9.0000	39.0000	0.0000	0.0000	0.0000
10.0000	45.0000	0.0000	0.0000	0.0000
11.0000	39.0000	0.0000	0.0000	0.0000
12.0000	35.0000	0.0000	0.0000	0.0000
13.0000	32.0000	0.0000	1.1875	13.0000
14.0000	39.0000	0.0000	8.3077	15.0000
15.0000	35.0000	0.0000	11.7429	17.0000
16.0000	34.0000	0.0000	11.8529	17.0000
17.0000	34.0000	0.0000	11.4706	22.0000
18.0000	38.0000	0.0000	10.2895	22.0000
19.0000	42.0000	0.0000	10.8333	26.0000
20.0000	35.0000	0.0000	11.8286	24.0000
21.0000	41.0000	0.0000	11.8780	38.0000
22.0000	37.0000	0.0000	4.2703	16.0000
23.0000	46.0000	0.0000	18.3913	31.0000
24.0000	36.0000	18.0000	29.4444	39.0000
25.0000	36.0000	0.0000	29.3056	40.0000
26.0000	48.0000	0.0000	24.9792	35.0000
27.0000	43.0000	0.0000	11.4651	26.0000
28.0000	40.0000	0.0000	5.6000	12.0000
29.0000	41.0000	0.0000	0.0000	0.0000
30.0000	38.0000	0.0000	0.0000	0.0000
31.0000	39.0000	0.0000	0.0000	0.0000

3 Error Tables

Data are for the month of August 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	2.4091	3.1319	0.5000	1.0000
2009.01	5.7963	5.1710	6.4215	1.3000	1.3000
2009.02	5.1088	4.5428	5.6749	0.7000	1.2000
2009.03	6.5648	6.3033	6.8262	0.3000	0.6000
2009.04	7.3838	7.1123	7.6552	0.4000	1.2000
2009.05	7.4883	7.1835	7.7931	1.6000	2.9000
2009.06	6.6562	6.3128	6.9995	3.2000	6.3000
2009.07	6.3365	6.0765	6.5965	3.6000	5.5000
2009.08	7.0316	6.7454	7.3177	0.0000	0.0000
2009.09	7.5796	7.3037	7.8556	4.5000	7.1000
2009.10	7.0695	6.6897	7.4492	4.5000	7.7000
2009.11	7.0246	6.8307	7.2186	3.3000	6.9000
2009.12	6.5379	6.3511	6.7247	10.4000	16.3000
2010.01	21.4034	18.9755	23.8314	13.3000	19.5000
2010.02	17.1481	14.8252	19.4710	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	18.3630	16.0914	20.6345	15.4000	24.0000
2010.04	20.3941	17.9969	22.7912	7.0000	10.4000
2010.05	24.4438	24.0106	24.8769	8.4000	8.7000
2010.06	20.3633	20.0244	20.7022	11.0000	13.6000
2010.07	21.2602	20.9518	21.5687	15.2000	16.1000
2010.08	22.7984	22.4238	23.1729	18.3000	19.6000
2010.09	25.5184	25.0974	25.9394	22.8000	25.2000
2010.10	24.0776	23.6616	24.4937	21.0000	23.5000
2010.11	24.5005	24.0556	24.9455	20.9000	21.6000
2010.12	21.8616	21.4210	22.3022	13.9000	14.5000
2011.01	76.2175	74.6412	77.7938	17.7000	18.7000
2011.02	66.3377	64.9216	67.7538	29.1000	29.6000
2011.03	69.1252	67.8064	70.4439	48.0000	55.8000
2011.04	77.6023	76.2084	78.9962	47.3000	54.4000
2011.05	78.8654	77.5327	80.1980	37.3000	41.5000
2011.06	65.7159	64.5640	66.8678	35.2000	37.0000
2011.07	67.8885	66.7258	69.0512	41.5000	43.8000
2011.08	73.6737	72.4865	74.8609	42.4000	50.5000
2011.09	81.0593	79.6582	82.4605	73.8000	78.0000
2011.10	76.5369	75.2491	77.8248	78.9000	88.0000
2011.11	77.6198	75.9983	79.2414	84.6000	96.7000
2011.12	68.1993	66.7947	69.6039	65.8000	73.0000
2012.01	81.8262	80.2199	83.4326	55.8000	58.2000
2012.02	70.0117	68.5895	71.4339	29.2000	33.1000
2012.03	73.6266	72.3236	74.9297	53.1000	64.1000
2012.04	81.4438	80.0212	82.8665	51.4000	55.2000
2012.05	84.3228	82.9332	85.7124	61.8000	69.0000
2012.06	69.5796	68.3973	70.7620	59.7000	64.5000
2012.07	72.2679	71.0826	73.4531	64.2000	51.3000
2012.08	75.6312	74.4173	76.8451	57.7000	63.1000
2012.09	83.7625	82.3095	85.2154	57.7000	61.5000
2012.10	79.9315	78.4697	81.3932	48.3000	53.3000
2012.11	81.0639	79.4387	82.6892	56.7000	61.4000
2012.12	71.3819	69.8340	72.9297	37.4000	40.8000
2013.01	90.9791	89.2424	92.7158	63.8000	62.9000
2013.02	77.9984	76.4305	79.5664	37.8000	38.0000
2013.03	79.4310	77.8200	81.0420	50.6000	57.9000
2013.04	88.8320	87.2739	90.3902	70.6000	72.4000
2013.05	89.9001	88.3030	91.4972	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	75.6815	74.3449	77.0181	51.0000	52.5000
2013.07	77.6271	76.3821	78.8721	57.0000	57.0000
2013.08	82.8019	81.4724	84.1313	60.0000	66.0000
2013.09	90.4072	88.7972	92.0171	34.6000	36.9000
2013.10	85.2124	83.6353	86.7895	74.5000	85.6000
2013.11	84.8424	82.9436	86.7412	73.9000	77.6000
2013.12	76.8387	75.2014	78.4759	77.8000	90.3000
2014.01	105.9875	103.7614	108.2136	77.4000	82.0000
2014.02	92.6810	90.8584	94.5036	93.9000	102.8000
2014.03	96.6086	94.8528	98.3644	80.9000	92.2000
2014.04	108.2014	106.3158	110.0870	76.9000	84.7000
2014.05	110.1812	108.3042	112.0583	72.3000	75.2000
2014.06	92.5445	90.9812	94.1078	67.2000	71.0000
2014.07	94.6376	93.0647	96.2105	72.5000	72.5000
2014.08	101.0954	99.5176	102.6732	71.2000	74.7000
2014.09	111.5283	109.5648	113.4919	83.2000	87.6000
2014.10	104.8093	102.8819	106.7368	59.5000	60.6000
2014.11	105.3895	103.1979	107.5811	65.8000	71.1000
2014.12	93.4874	91.3236	95.6511	75.8000	78.0000
2015.01	65.5552	64.2442	66.8661	65.9000	67.0000
2015.02	55.9093	54.6784	57.1402	42.4000	44.8000
2015.03	59.0372	57.9566	60.1178	38.0000	38.4000
2015.04	65.6953	64.5169	66.8738	49.0000	54.4000
2015.05	67.1933	66.0800	68.3066	56.3000	58.8000
2015.06	56.3488	55.3314	57.3663	50.2000	68.3000
2015.07	57.1137	56.1376	58.0899	47.9000	65.8000
2015.08	62.2594	61.2116	63.3072	39.5000	57.2000
2015.09	67.7901	66.5569	69.0233	49.2000	72.1000
2015.10	64.2049	62.9639	65.4460	39.3000	48.3000
2015.11	65.2843	63.8662	66.7025	39.6000	55.9000
2015.12	57.9834	56.7092	59.2576	36.4000	44.8000
2016.01	35.8772	35.1284	36.6261	33.7000	43.3000
2016.02	30.7437	30.1028	31.3846	38.3000	46.8000
2016.03	31.9428	31.3053	32.5802	30.5000	38.9000
2016.04	35.4377	34.7616	36.1138	26.6000	30.9000
2016.05	36.3397	35.6742	37.0052	33.7000	48.4000
2016.06	30.1400	29.6250	30.6550	13.1000	19.5000
2016.07	31.1193	30.6179	31.6207	21.2000	27.5000
2016.08	33.5536	32.9623	34.1449	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	37.3982	36.7143	38.0821	27.7000	37.1000
2016.10	35.0183	34.3419	35.6948	22.7000	31.7000
2016.11	35.2047	34.4682	35.9412	14.0000	22.2000
2016.12	31.6685	30.9923	32.3447	11.1000	20.0000
2017.01	19.5181	19.1076	19.9286	18.4000	26.2000
2017.02	16.7900	16.4215	17.1586	14.4000	20.6000
2017.03	17.5991	17.2639	17.9343	11.3000	15.5000
2017.04	19.7404	19.3918	20.0889	21.6000	33.2000
2017.05	19.9146	19.5704	20.2588	12.5000	18.1000
2017.06	16.5334	16.2572	16.8096	15.5000	19.3000
2017.07	17.1434	16.8705	17.4162	11.5000	16.3000
2017.08	18.4163	18.0995	18.7331	22.8000	35.7000
2017.09	20.7866	20.3826	21.1906	34.6000	42.9000
2017.10	19.0501	18.6718	19.4284	10.5000	11.0000
2017.11	19.0073	18.6109	19.4036	4.2000	5.6000
2017.12	16.9818	16.7229	17.2407	4.0000	4.6000
2018.01	6.4764	6.3374	6.6155	3.1000	6.3000
2018.02	5.5171	5.3842	5.6499	6.8000	11.8000
2018.03	5.6831	5.5685	5.7977	1.1000	1.2000
2018.04	6.3111	6.1845	6.4376	4.7000	7.5000
2018.05	6.5175	6.3905	6.6445	8.4000	14.0000
2018.06	5.3944	5.2945	5.4944	10.2000	13.6000
2018.07	5.5675	5.5021	5.6329	0.5000	1.7000
2018.08	5.9231	5.8144	6.0318	5.9000	9.5000

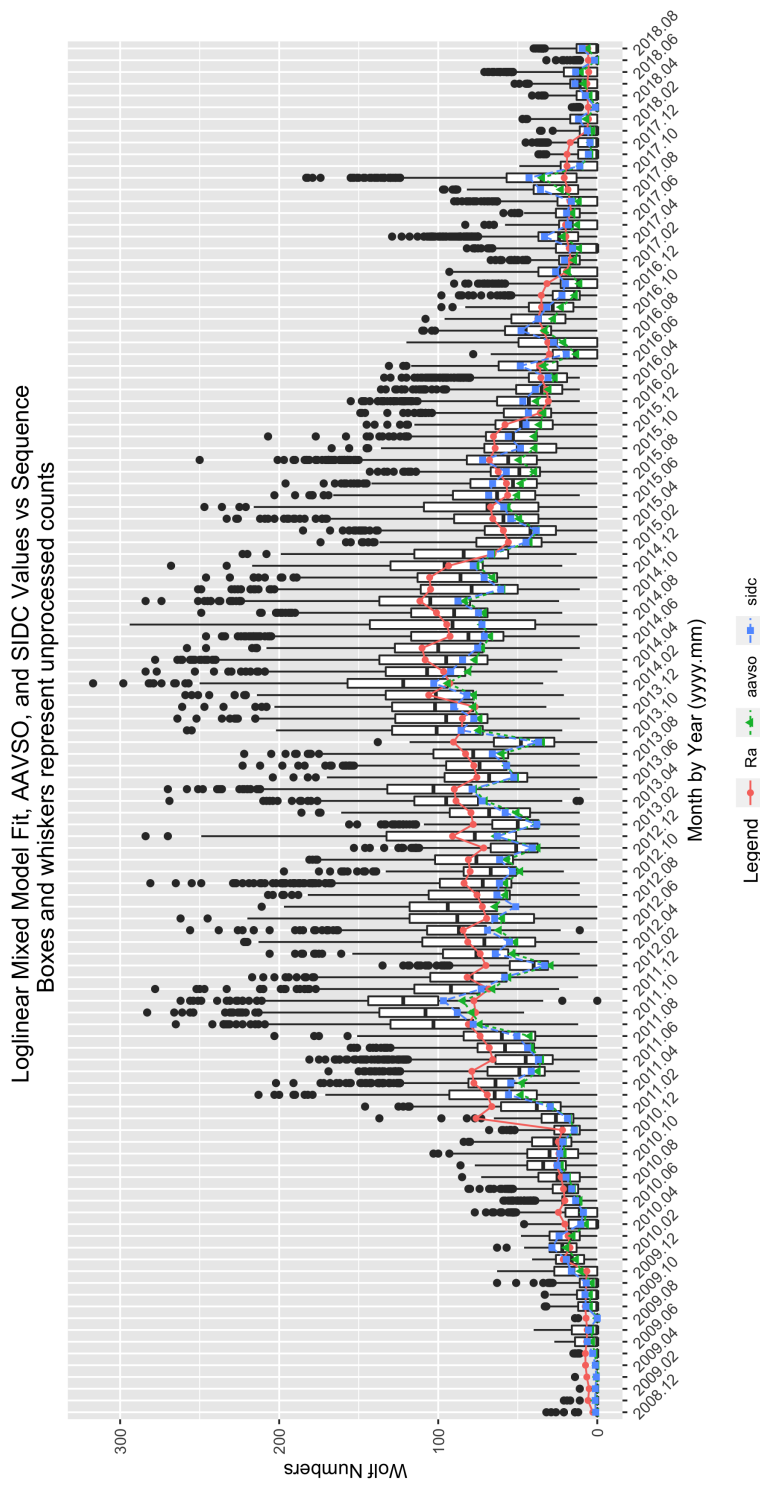


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201808 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4226	0.3181	4.4725	0.0000
seeF	-0.2184	0.0060	-36.2549	0.0000
seeG	-0.1155	0.0052	-22.0084	0.0000
seeM	-0.1951	0.0247	-7.8987	0.0000
seeP	-0.3249	0.0086	-37.6424	0.0000
sidc1	0.1510	0.0690	2.1867	0.0288
year2009	0.6444	0.3191	2.0197	0.0434
year2010	1.8551	0.3169	5.8545	0.0000
year2011	2.9762	0.3168	9.3956	0.0000
year2012	3.0139	0.3168	9.5147	0.0000
year2013	3.1100	0.3168	9.8181	0.0000
year2014	3.3070	0.3168	10.4401	0.0000
year2015	2.8225	0.3168	8.9101	0.0000
year2016	2.2055	0.3168	6.9617	0.0000
year2017	1.6006	0.3168	5.0516	0.0000
year2018	0.4885	0.3172	1.5397	0.1236
mon2	-0.1450	0.0095	-15.2642	0.0000
mon3	-0.1113	0.0089	-12.4599	0.0000
mon4	-0.0088	0.0086	-1.0231	0.3063
mon5	0.0035	0.0085	0.4083	0.6830
mon6	-0.1811	0.0088	-20.4720	0.0000
mon7	-0.1572	0.0086	-18.2752	0.0000
mon8	-0.0828	0.0084	-9.8379	0.0000
mon9	0.0235	0.0085	2.7728	0.0056
mon10	-0.0325	0.0087	-3.7379	0.0002
mon11	-0.0122	0.0091	-1.3400	0.1802
mon12	-0.1220	0.0093	-13.1701	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: 201808 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 0.00	Length:103944	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 : 6.587	Mean :15.73		Mean :0.2629
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: 201808 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:103944	Length:103944
1st Qu.: 1.000	1st Qu.: 3.00	1st Qu.: 16.00	Class :character	Class :character
Median : 3.000	Median : 14.00	Median : 46.00	Mode :character	Mode :character
Mean : 3.436	Mean : 20.57	Mean : 54.93		
3rd Qu.: 5.000	3rd Qu.: 30.00	3rd Qu.: 83.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201808 Summary of Sunspot Numbers

inst	filter	unit
Length:103944	Length:103944	Length:103944
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201808 Summary of Sunspot Numbers

aperture	eyep	foclen	mag
Min. : 0.0	Min. : 0.0	Min. : 0	Min. : 0.0
1st Qu.: 76.0	1st Qu.: 2.0	1st Qu.: 750	1st Qu.: 40.0
Median : 90.0	Median : 13.0	Median :1000	Median : 57.5
Mean : 109.1	Mean : 20.4	Mean :1094	Mean : 186.5
3rd Qu.: 125.0	3rd Qu.: 23.0	3rd Qu.:1260	3rd Qu.: 76.0
Max. :1524.0	Max. :2010.0	Max. :4300	Max. :4591.0

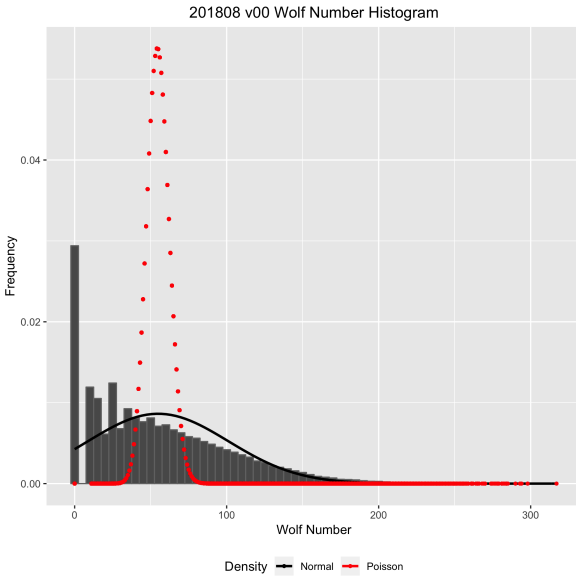


Figure 3: Box plots of raw Wolf number (w) by observer rank.

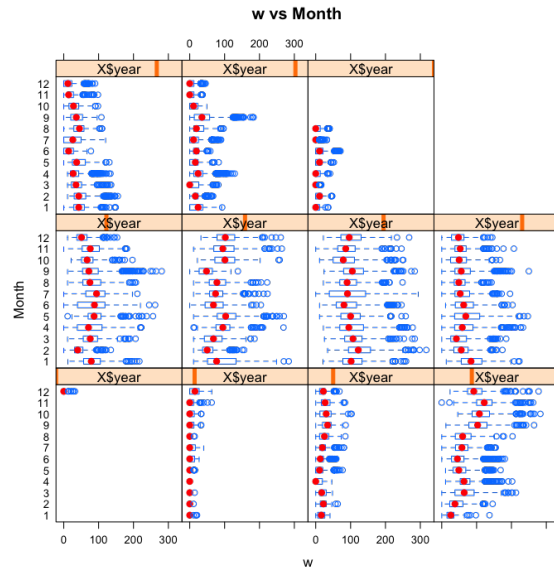


Figure 4: Box plots of raw Wolf number (w) by month and year.

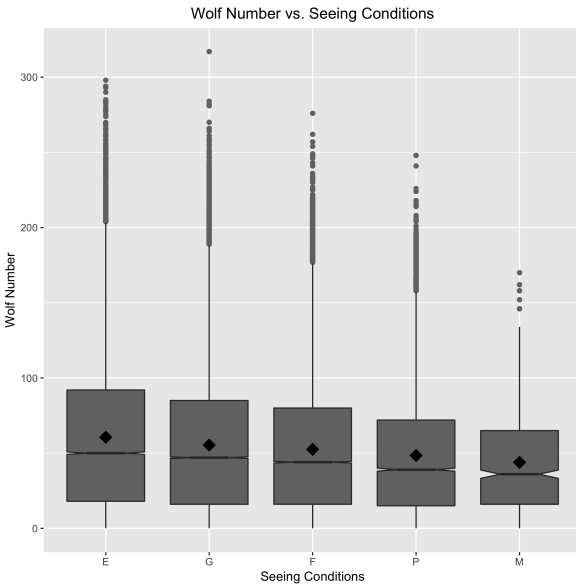


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

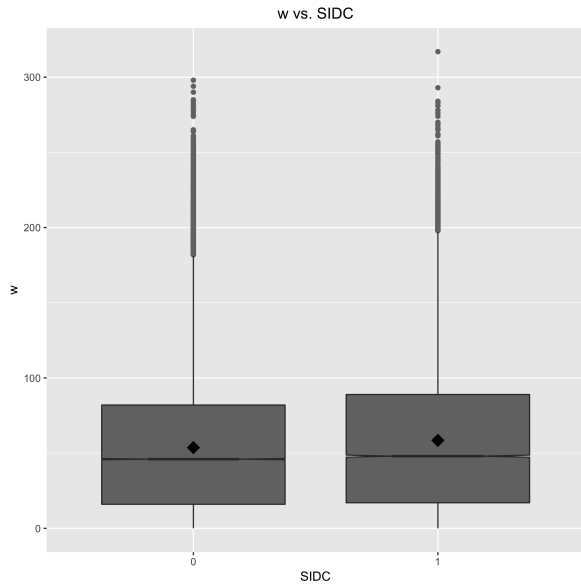


Figure 6: Box plots of raw Wolf number (w) by organization.

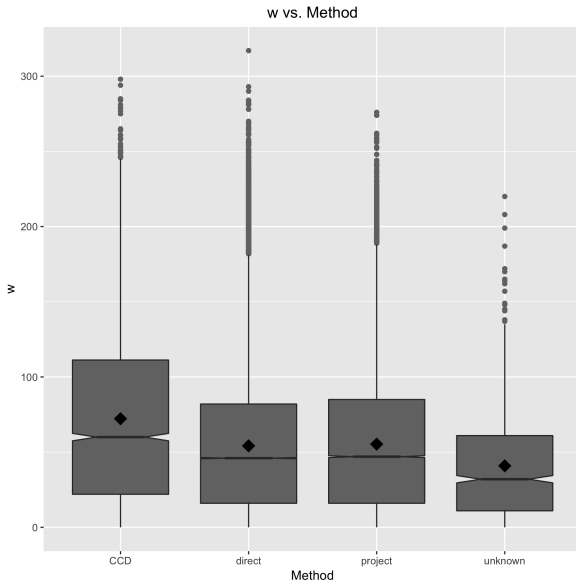


Figure 7: Box plots of raw Wolf number (w) by observer rank.

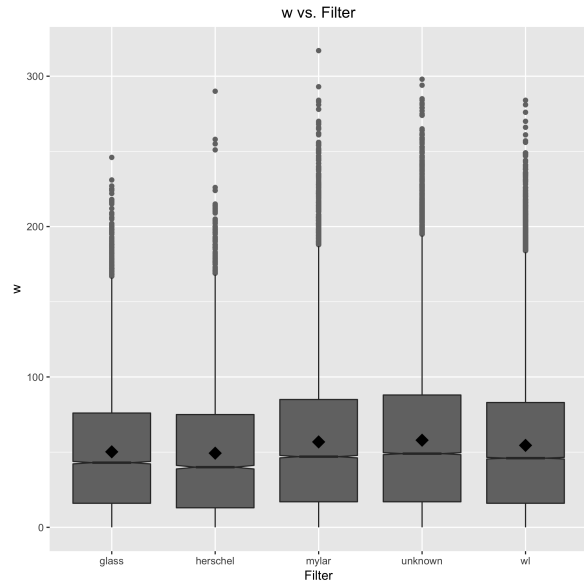


Figure 8: Box plots of raw Wolf number (w) by month and year.

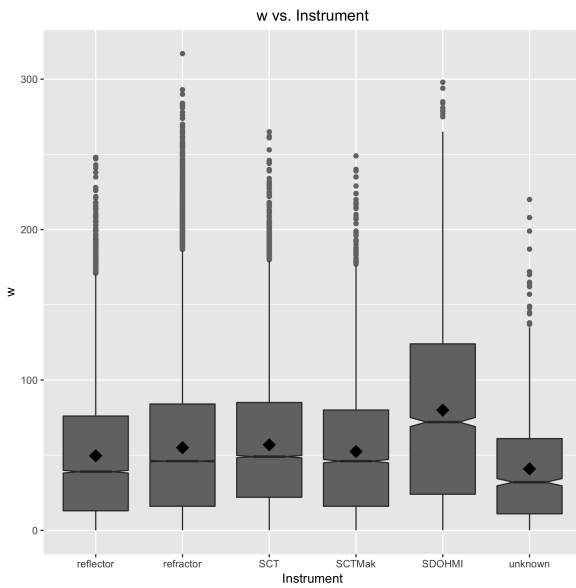


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

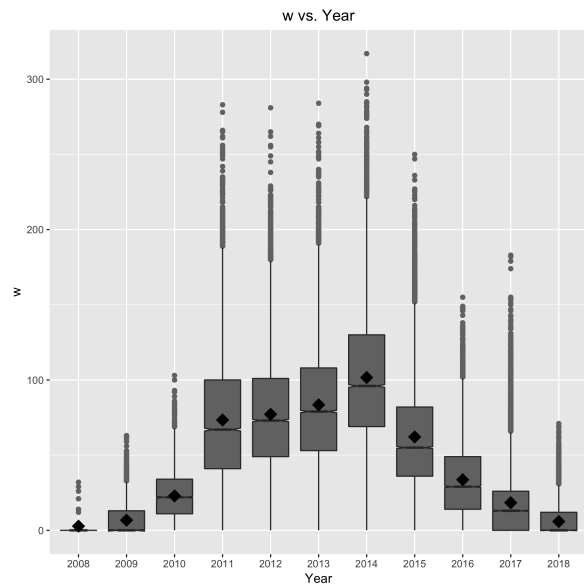


Figure 10: Box plots of raw Wolf number (w) by organization.

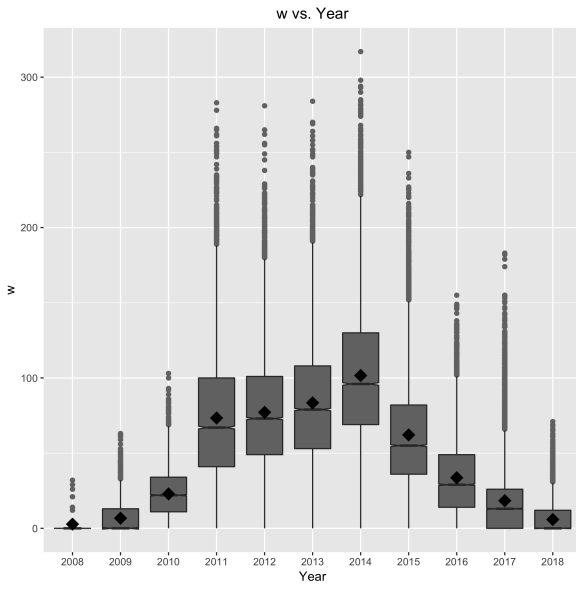


Figure 11: Box plots of raw Wolf number (w) by year.

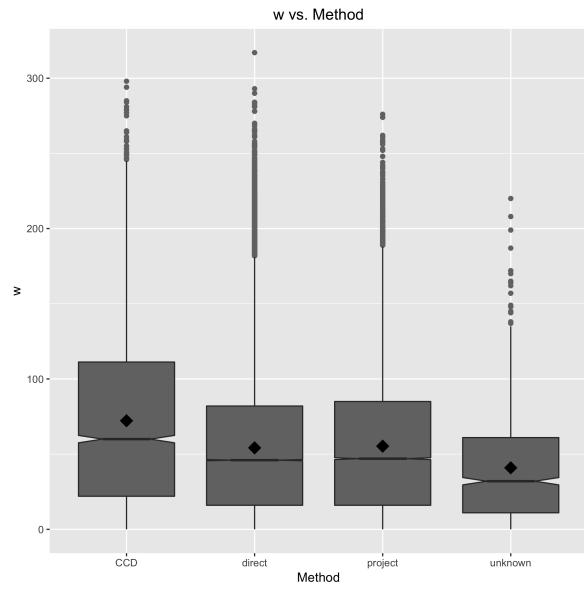


Figure 12: Box plots of raw Wolf number (w) by observing method.