

Monthly Report (00)
201806 Data Set

Wednesday 11th July, 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.

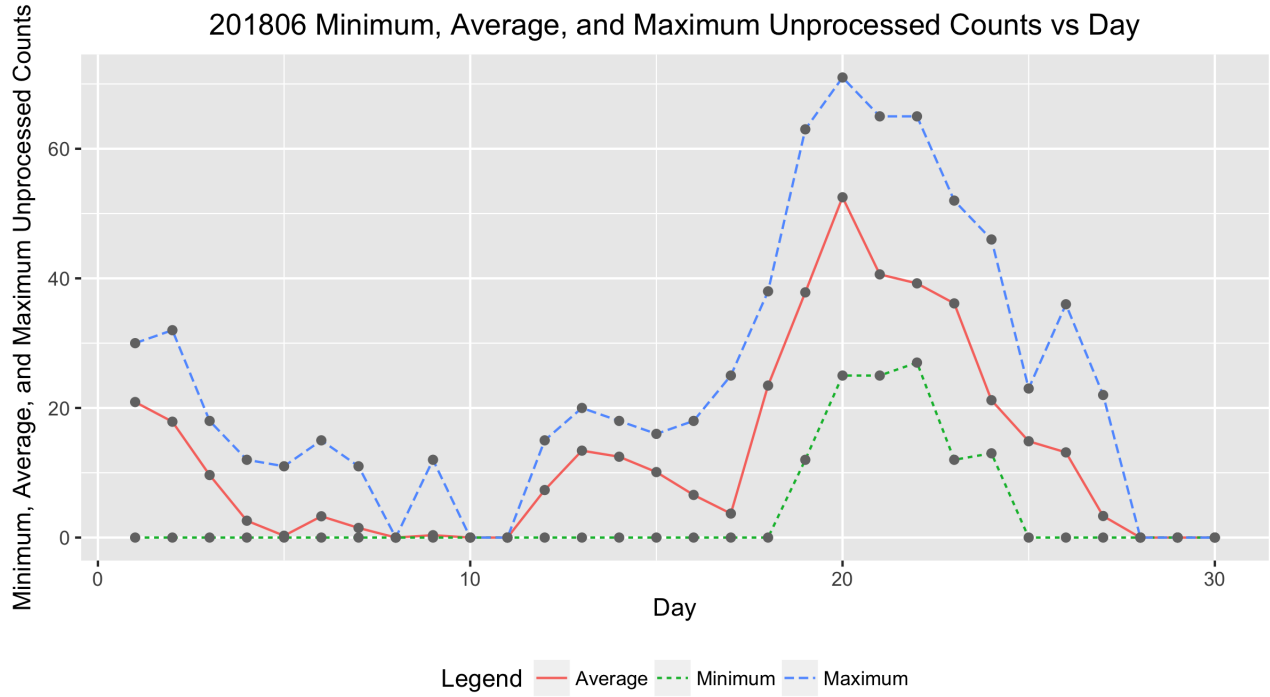


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201806 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	35.0000	0.0000	20.9143	30.0000
2.0000	41.0000	0.0000	17.8780	32.0000
3.0000	34.0000	0.0000	9.6471	18.0000
4.0000	30.0000	0.0000	2.6000	12.0000
5.0000	39.0000	0.0000	0.2821	11.0000
6.0000	39.0000	0.0000	3.2821	15.0000
7.0000	37.0000	0.0000	1.4865	11.0000
8.0000	40.0000	0.0000	0.0000	0.0000
9.0000	33.0000	0.0000	0.3636	12.0000
10.0000	36.0000	0.0000	0.0000	0.0000
11.0000	37.0000	0.0000	0.0000	0.0000
12.0000	39.0000	0.0000	7.3333	15.0000
13.0000	34.0000	0.0000	13.4118	20.0000
14.0000	40.0000	0.0000	12.4750	18.0000
15.0000	38.0000	0.0000	10.1053	16.0000
16.0000	38.0000	0.0000	6.5789	18.0000
17.0000	37.0000	0.0000	3.7027	25.0000
18.0000	35.0000	0.0000	23.4571	38.0000
19.0000	37.0000	12.0000	37.8378	63.0000
20.0000	37.0000	25.0000	52.5135	71.0000
21.0000	49.0000	25.0000	40.6122	65.0000
22.0000	44.0000	27.0000	39.2273	65.0000
23.0000	37.0000	12.0000	36.1351	52.0000
24.0000	40.0000	13.0000	21.2000	46.0000
25.0000	44.0000	0.0000	14.8636	23.0000
26.0000	42.0000	0.0000	13.1429	36.0000
27.0000	37.0000	0.0000	3.3243	22.0000
28.0000	42.0000	0.0000	0.0000	0.0000
29.0000	43.0000	0.0000	0.0000	0.0000
30.0000	44.0000	0.0000	0.0000	0.0000

3 Error Tables

Data are for the month of June 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	2.4116	3.1293	0.5000	1.0000
2009.01	5.7816	5.1626	6.4005	1.3000	1.3000
2009.02	5.0937	4.5336	5.6538	0.7000	1.2000
2009.03	6.5359	6.2741	6.7977	0.3000	0.6000
2009.04	7.3526	7.0810	7.6241	0.4000	1.2000
2009.05	7.4547	7.1497	7.7597	1.6000	2.9000
2009.06	6.6265	6.2832	6.9698	3.2000	6.3000
2009.07	6.4911	6.2235	6.7586	3.6000	5.5000
2009.08	7.0058	6.7195	7.2921	0.0000	0.0000
2009.09	7.5772	7.3002	7.8543	4.5000	7.1000
2009.10	7.0680	6.6871	7.4489	4.5000	7.7000
2009.11	7.0258	6.8325	7.2190	3.3000	6.9000
2009.12	6.5389	6.3527	6.7252	10.4000	16.3000
2010.01	21.2990	18.9002	23.6978	13.3000	19.5000
2010.02	17.0731	14.7785	19.3678	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	18.2832	16.0381	20.5284	15.4000	24.0000
2010.04	20.3080	17.9384	22.6777	7.0000	10.4000
2010.05	24.2963	23.8671	24.7254	8.4000	8.7000
2010.06	20.2372	19.9008	20.5736	11.0000	13.6000
2010.07	21.7390	21.4239	22.0540	15.2000	16.1000
2010.08	22.6737	22.3015	23.0458	18.3000	19.6000
2010.09	25.4667	25.0470	25.8863	22.8000	25.2000
2010.10	24.0267	23.6122	24.4412	21.0000	23.5000
2010.11	24.4525	24.0091	24.8959	20.9000	21.6000
2010.12	21.8165	21.3777	22.2553	13.9000	14.5000
2011.01	75.9085	74.3397	77.4773	17.7000	18.7000
2011.02	66.0501	64.6398	67.4605	29.1000	29.6000
2011.03	68.8563	67.5448	70.1678	48.0000	55.8000
2011.04	77.3065	75.9186	78.6944	47.3000	54.4000
2011.05	78.5426	77.2168	79.8684	37.3000	41.5000
2011.06	65.4347	64.2882	66.5813	35.2000	37.0000
2011.07	69.5553	68.3641	70.7465	41.5000	43.8000
2011.08	73.4178	72.2351	74.6005	42.4000	50.5000
2011.09	81.0630	79.6628	82.4633	73.8000	78.0000
2011.10	76.5329	75.2462	77.8196	78.9000	88.0000
2011.11	77.6198	75.9999	79.2397	84.6000	96.7000
2011.12	68.1975	66.7936	69.6013	65.8000	73.0000
2012.01	81.5099	79.9111	83.1088	55.8000	58.2000
2012.02	69.7290	68.3133	71.1447	29.2000	33.1000
2012.03	73.3486	72.0515	74.6458	53.1000	64.1000
2012.04	81.1425	79.7260	82.5589	51.4000	55.2000
2012.05	83.9815	82.5982	85.3648	61.8000	69.0000
2012.06	69.3024	68.1259	70.4789	59.7000	64.5000
2012.07	74.0636	72.8494	75.2777	64.2000	51.3000
2012.08	75.3833	74.1739	76.5926	57.7000	63.1000
2012.09	83.7814	82.3288	85.2339	57.7000	61.5000
2012.10	79.9476	78.4865	81.4088	48.3000	53.3000
2012.11	81.0755	79.4514	82.6996	56.7000	61.4000
2012.12	71.3908	69.8439	72.9377	37.4000	40.8000
2013.01	90.6158	88.8874	92.3443	63.8000	62.9000
2013.02	77.6688	76.1085	79.2290	37.8000	38.0000
2013.03	79.1148	77.5113	80.7183	50.6000	57.9000
2013.04	88.4971	86.9464	90.0477	70.6000	72.4000
2013.05	89.5300	87.9405	91.1194	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	75.3652	74.0351	76.6952	51.0000	52.5000
2013.07	79.5494	78.2743	80.8245	57.0000	57.0000
2013.08	82.5215	81.1973	83.8456	60.0000	66.0000
2013.09	90.4202	88.8114	92.0290	34.6000	36.9000
2013.10	85.2141	83.6386	86.7896	74.5000	85.6000
2013.11	84.8485	82.9513	86.7458	73.9000	77.6000
2013.12	76.8405	75.2048	78.4762	77.8000	90.3000
2014.01	105.5771	103.3611	107.7931	77.4000	82.0000
2014.02	92.2975	90.4836	94.1113	93.9000	102.8000
2014.03	96.2451	94.4976	97.9926	80.9000	92.2000
2014.04	107.8112	105.9336	109.6887	76.9000	84.7000
2014.05	109.7485	107.8807	111.6162	72.3000	75.2000
2014.06	92.1773	90.6217	93.7328	67.2000	71.0000
2014.07	96.9992	95.3883	98.6101	72.5000	72.5000
2014.08	100.7640	99.1926	102.3355	71.2000	74.7000
2014.09	111.5519	109.5892	113.5146	83.2000	87.6000
2014.10	104.8233	102.8966	106.7499	59.5000	60.6000
2014.11	105.3994	103.2090	107.5898	65.8000	71.1000
2014.12	93.4931	91.3310	95.6551	75.8000	78.0000
2015.01	65.3169	64.0119	66.6219	65.9000	67.0000
2015.02	55.6861	54.4606	56.9115	42.4000	44.8000
2015.03	58.8103	57.7343	59.8864	38.0000	38.4000
2015.04	65.4537	64.2789	66.6285	49.0000	54.4000
2015.05	66.9237	65.8152	68.0322	56.3000	58.8000
2015.06	56.1229	55.1091	57.1366	50.2000	68.3000
2015.07	58.5250	57.5241	59.5258	47.9000	65.8000
2015.08	62.0594	61.0145	63.1042	39.5000	57.2000
2015.09	67.8023	66.5688	69.0358	49.2000	72.1000
2015.10	64.2143	62.9734	65.4553	39.3000	48.3000
2015.11	65.2980	63.8805	66.7156	39.6000	55.9000
2015.12	57.9948	56.7213	59.2683	36.4000	44.8000
2016.01	35.7318	34.9863	36.4772	33.7000	43.3000
2016.02	30.6114	29.9736	31.2492	38.3000	46.8000
2016.03	31.8157	31.1812	32.4503	30.5000	38.9000
2016.04	35.2946	34.6212	35.9680	26.6000	30.9000
2016.05	36.1884	35.5259	36.8509	33.7000	48.4000
2016.06	30.0125	29.4997	30.5252	13.1000	19.5000
2016.07	31.8938	31.3799	32.4076	21.2000	27.5000
2016.08	33.4399	32.8507	34.0291	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	37.3977	36.7139	38.0815	27.7000	37.1000
2016.10	35.0199	34.3438	35.6960	22.7000	31.7000
2016.11	35.2020	34.4659	35.9381	14.0000	22.2000
2016.12	31.6671	30.9910	32.3432	11.1000	20.0000
2017.01	19.4422	19.0334	19.8511	18.4000	26.2000
2017.02	16.7224	16.3555	17.0894	14.4000	20.6000
2017.03	17.5325	17.1986	17.8664	11.3000	15.5000
2017.04	19.6686	19.3214	20.0159	21.6000	33.2000
2017.05	19.8358	19.4930	20.1785	12.5000	18.1000
2017.06	16.4687	16.1937	16.7436	15.5000	19.3000
2017.07	17.5706	17.2911	17.8501	11.5000	16.3000
2017.08	18.3519	18.0376	18.6663	22.8000	35.7000
2017.09	20.7551	20.3616	21.1487	34.6000	42.9000
2017.10	19.0477	18.6727	19.4228	10.5000	11.0000
2017.11	19.0058	18.6101	19.4014	4.2000	5.6000
2017.12	16.9854	16.7273	17.2436	4.0000	4.6000
2018.01	8.4062	8.2280	8.5843	3.1000	6.3000
2018.02	7.1723	7.0039	7.3408	6.8000	11.8000
2018.03	7.4055	7.2575	7.5535	1.1000	1.2000
2018.04	8.2112	8.0507	8.3716	4.7000	7.5000
2018.05	8.4709	8.3124	8.6294	8.4000	14.0000
2018.06	7.0481	6.9207	7.1755	10.2000	13.6000

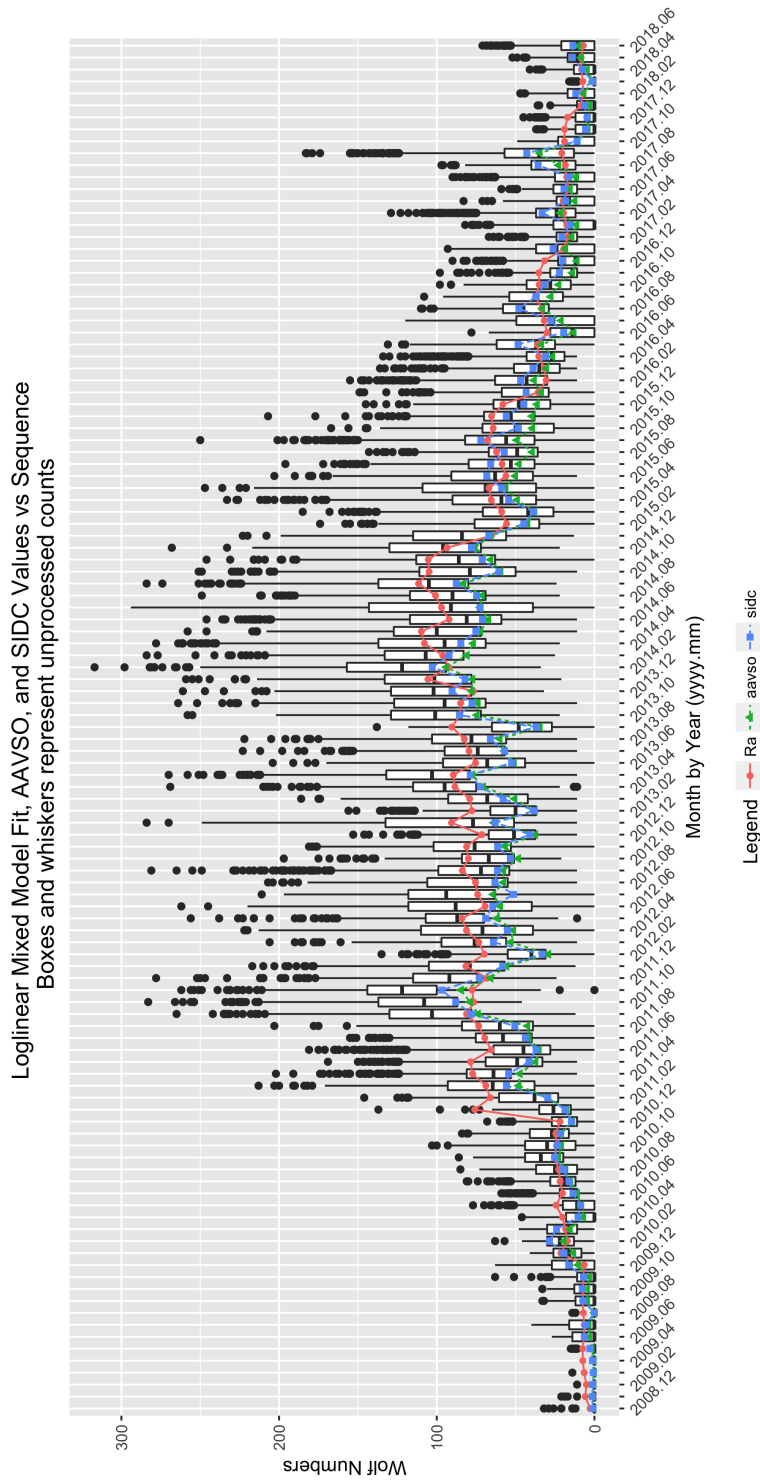


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201806 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4273	0.3196	4.4658	0.0000
seeF	-0.2183	0.0061	-36.0324	0.0000
seeG	-0.1154	0.0053	-21.8701	0.0000
seeM	-0.1935	0.0248	-7.7911	0.0000
seeP	-0.3252	0.0087	-37.4654	0.0000
sidc1	0.1198	0.0671	1.7836	0.0745
year2009	0.6460	0.3207	2.0146	0.0439
year2010	1.8547	0.3184	5.8245	0.0000
year2011	2.9779	0.3183	9.3548	0.0000
year2012	3.0158	0.3183	9.4741	0.0000
year2013	3.1118	0.3183	9.7756	0.0000
year2014	3.3089	0.3183	10.3949	0.0000
year2015	2.8244	0.3183	8.8725	0.0000
year2016	2.2071	0.3184	6.9327	0.0000
year2017	1.6024	0.3184	5.0328	0.0000
year2018	0.7555	0.3189	2.3689	0.0178
mon2	-0.1452	0.0095	-15.2073	0.0000
mon3	-0.1112	0.0090	-12.3938	0.0000
mon4	-0.0086	0.0087	-0.9938	0.3203
mon5	0.0033	0.0085	0.3922	0.6949
mon6	-0.1812	0.0089	-20.3799	0.0000
mon7	-0.1286	0.0087	-14.8700	0.0000
mon8	-0.0821	0.0085	-9.6721	0.0000
mon9	0.0276	0.0085	3.2463	0.0012
mon10	-0.0285	0.0087	-3.2542	0.0011
mon11	-0.0082	0.0091	-0.8936	0.3715
mon12	-0.1180	0.0093	-12.6737	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: 201806 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 1.00	Length:100197	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 : 6.559	Mean :15.73		Mean :0.2632
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: 201806 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:100197	Length:100197
1st Qu.: 1.000	1st Qu.: 4.00	1st Qu.: 21.00	Class :character	Class :character
Median : 3.000	Median : 14.00	Median : 48.00	Mode :character	Mode :character
Mean : 3.557	Mean : 21.31	Mean : 56.88		
3rd Qu.: 5.000	3rd Qu.: 31.00	3rd Qu.: 85.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201806 Summary of Sunspot Numbers

inst	filter	unit
Length:100197	Length:100197	Length:100197
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201806 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 : 110.3	Mean : 18.71	Mean :1109	Mean : 187.2
3rd Qu.: 125.0	3rd Qu.: 23.00	3rd Qu.:1296	3rd Qu.: 76.0
Max. :1524.0	Max. :2010.00	Max. :4300	Max. :4591.0

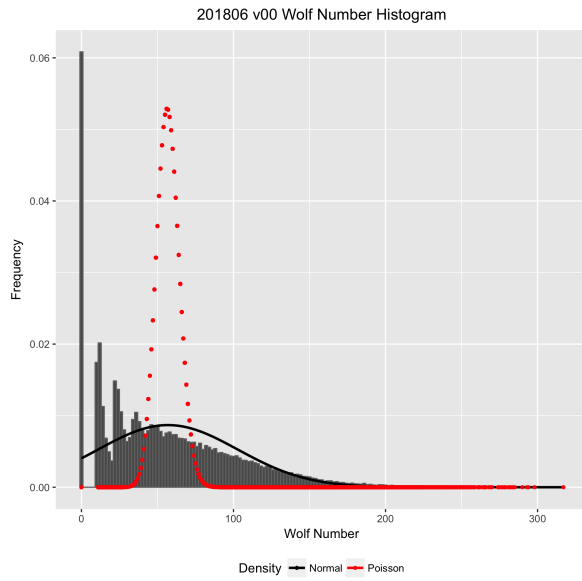


Figure 3: Box plots of raw Wolf number (w) by observer rank.

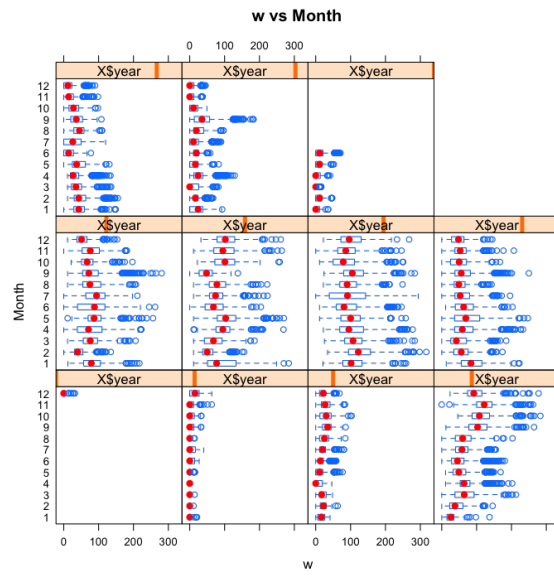


Figure 4: Box plots of raw Wolf number (w) by month and year.

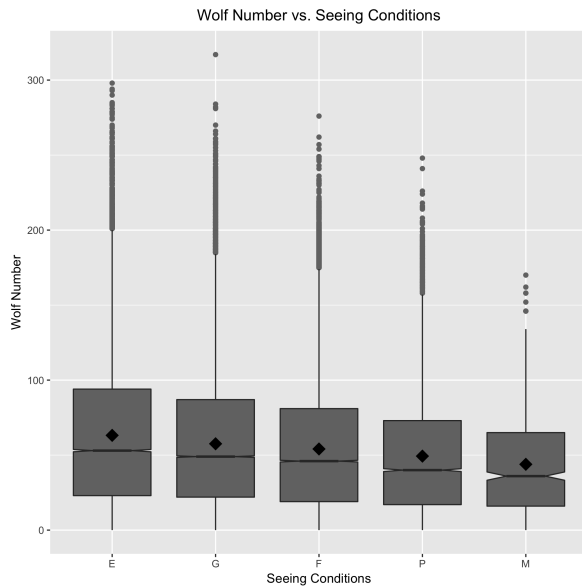


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

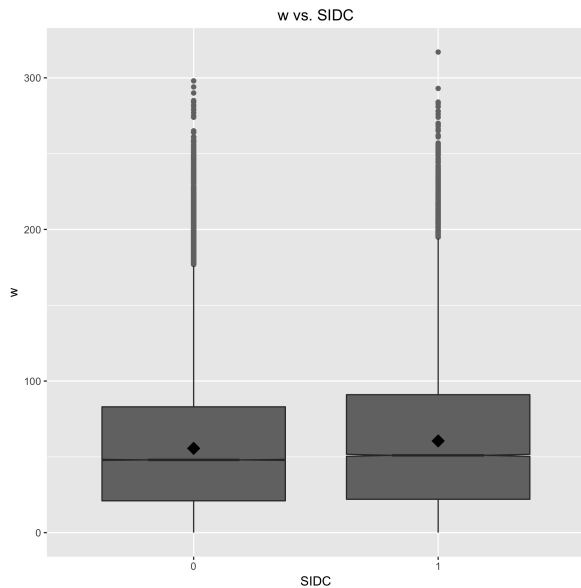


Figure 6: Box plots of raw Wolf number (w) by organization.

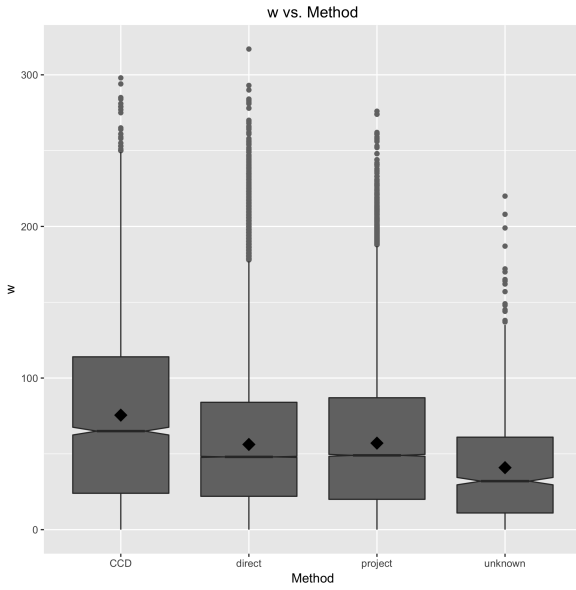


Figure 7: Box plots of raw Wolf number (w) by observer rank.

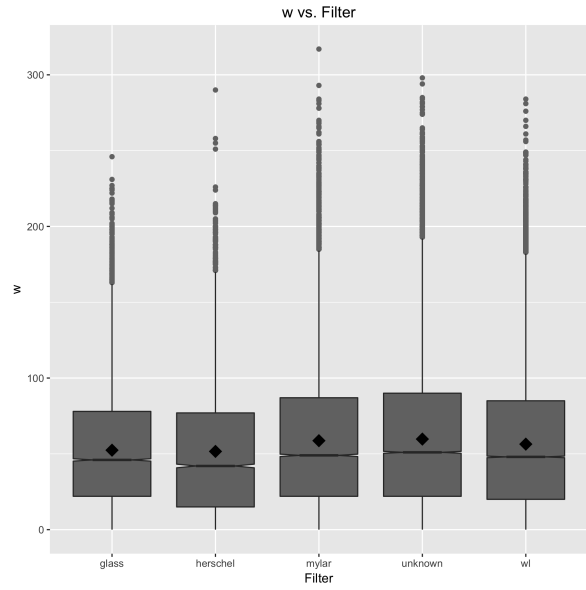


Figure 8: Box plots of raw Wolf number (w) by month and year.

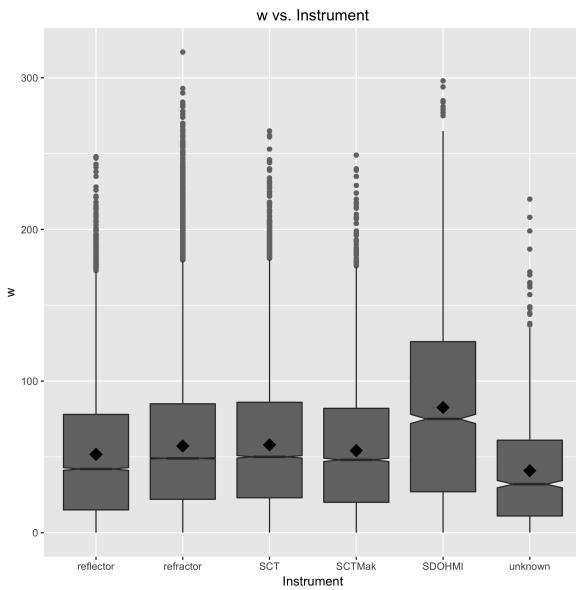


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

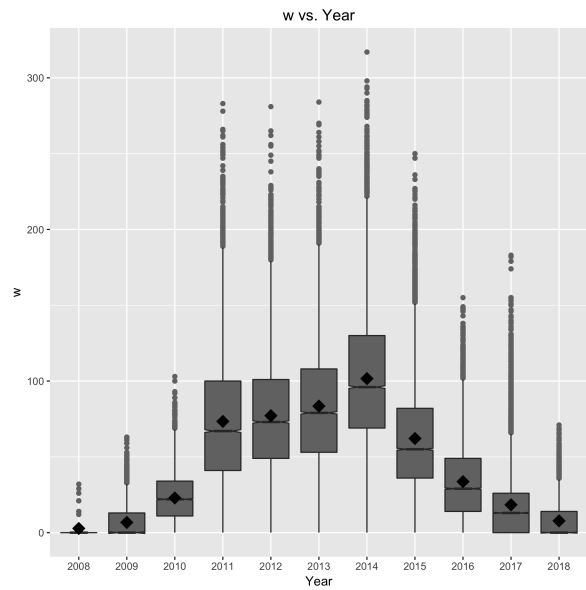


Figure 10: Box plots of raw Wolf number (w) by organization.

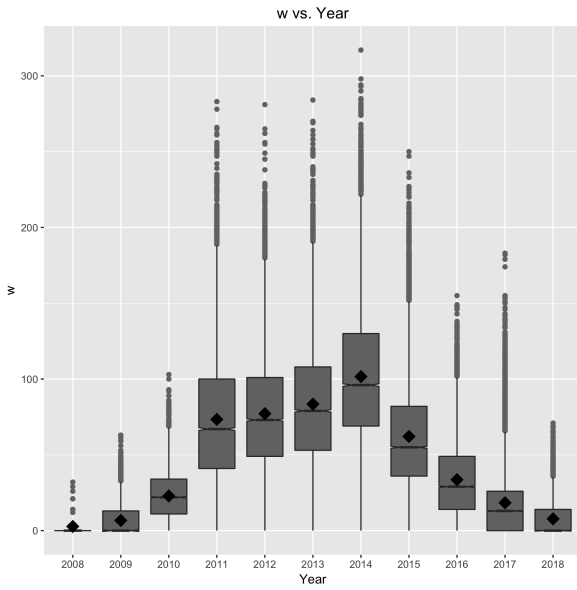


Figure 11: Box plots of raw Wolf number (w) by year.

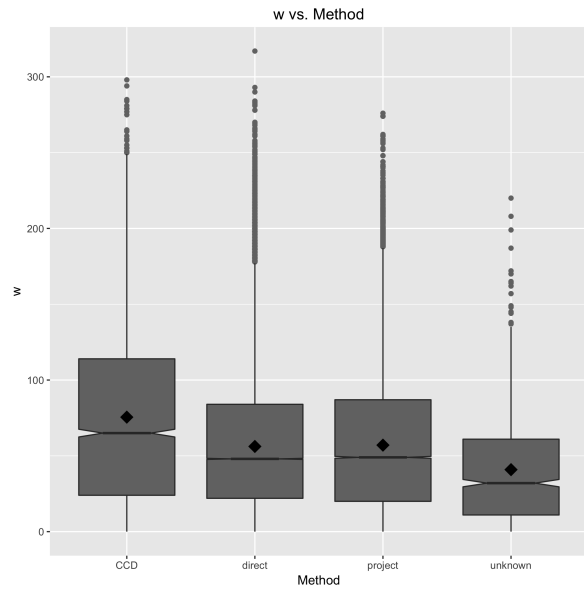


Figure 12: Box plots of raw Wolf number (w) by observing method.