

Monthly Report (00)

201803 Data Set

Friday 13th April, 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.

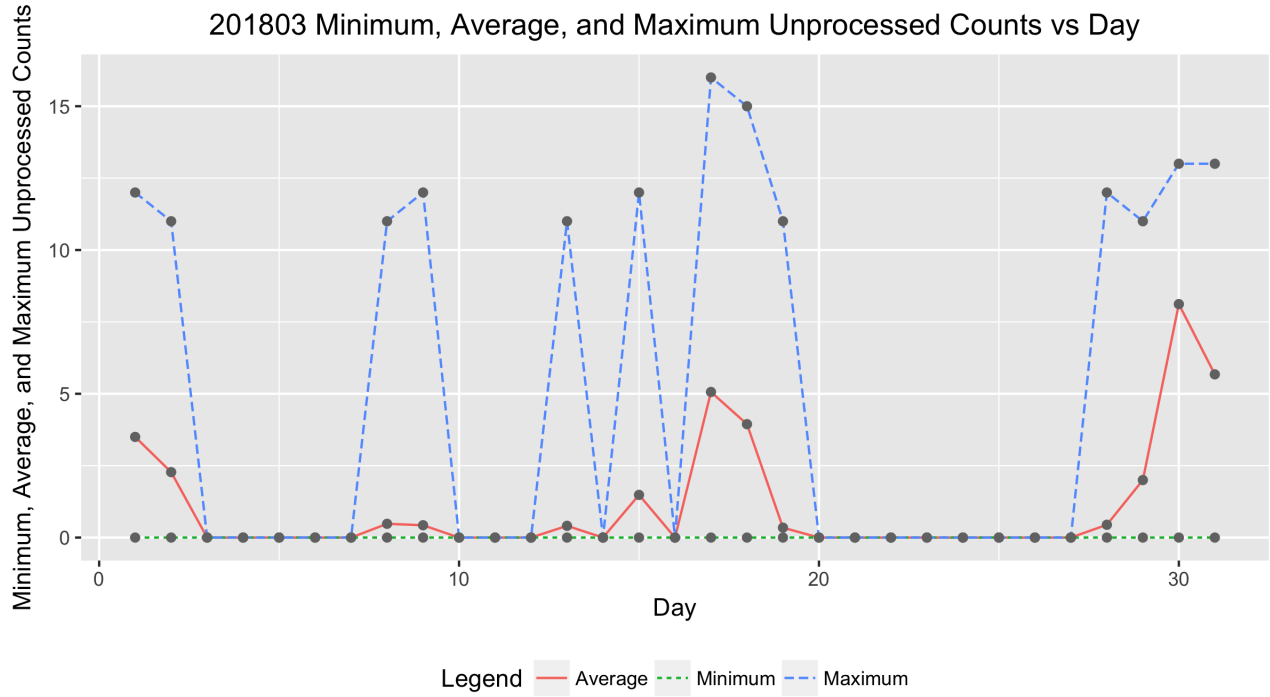


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201803 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	26.0000	0.0000	3.5000	12.0000
2.0000	29.0000	0.0000	2.2759	11.0000
3.0000	36.0000	0.0000	0.0000	0.0000
4.0000	37.0000	0.0000	0.0000	0.0000
5.0000	31.0000	0.0000	0.0000	0.0000
6.0000	34.0000	0.0000	0.0000	0.0000
7.0000	22.0000	0.0000	0.0000	0.0000
8.0000	23.0000	0.0000	0.4783	11.0000
9.0000	28.0000	0.0000	0.4286	12.0000
10.0000	32.0000	0.0000	0.0000	0.0000
11.0000	34.0000	0.0000	0.0000	0.0000
12.0000	29.0000	0.0000	0.0000	0.0000
13.0000	27.0000	0.0000	0.4074	11.0000
14.0000	28.0000	0.0000	0.0000	0.0000
15.0000	31.0000	0.0000	1.4839	12.0000
16.0000	27.0000	0.0000	0.0000	0.0000
17.0000	32.0000	0.0000	5.0625	16.0000
18.0000	35.0000	0.0000	3.9429	15.0000
19.0000	32.0000	0.0000	0.3438	11.0000
20.0000	29.0000	0.0000	0.0000	0.0000
21.0000	30.0000	0.0000	0.0000	0.0000
22.0000	31.0000	0.0000	0.0000	0.0000
23.0000	30.0000	0.0000	0.0000	0.0000
24.0000	35.0000	0.0000	0.0000	0.0000
25.0000	40.0000	0.0000	0.0000	0.0000
26.0000	36.0000	0.0000	0.0000	0.0000
27.0000	30.0000	0.0000	0.0000	0.0000
28.0000	27.0000	0.0000	0.4444	12.0000
29.0000	33.0000	0.0000	2.0000	11.0000
30.0000	34.0000	0.0000	8.1176	13.0000
31.0000	37.0000	0.0000	5.6757	13.0000

3 Error Tables

Data are for the month of March 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	2.4100	3.1310	0.5000	1.0000
2009.01	5.8113	5.1863	6.4364	1.3000	1.3000
2009.02	5.1224	4.5565	5.6884	0.7000	1.2000
2009.03	6.5779	6.3154	6.8403	0.3000	0.6000
2009.04	7.3835	7.1115	7.6556	0.4000	1.2000
2009.05	7.4187	7.1163	7.7211	1.6000	2.9000
2009.06	6.5174	6.1810	6.8537	3.2000	6.3000
2009.07	6.4893	6.2228	6.7557	3.6000	5.5000
2009.08	7.0018	6.7164	7.2873	0.0000	0.0000
2009.09	7.5767	7.3003	7.8530	4.5000	7.1000
2009.10	7.0659	6.6860	7.4459	4.5000	7.7000
2009.11	7.0247	6.8309	7.2185	3.3000	6.9000
2009.12	6.5411	6.3542	6.7279	10.4000	16.3000
2010.01	21.4553	19.0276	23.8831	13.3000	19.5000
2010.02	17.2008	14.8773	19.5244	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	18.4082	16.1375	20.6790	15.4000	24.0000
2010.04	20.4078	18.0156	22.8000	7.0000	10.4000
2010.05	24.2392	23.8107	24.6677	8.4000	8.7000
2010.06	19.9565	19.6249	20.2881	11.0000	13.6000
2010.07	21.7909	21.4755	22.1064	15.2000	16.1000
2010.08	22.7280	22.3550	23.1010	18.3000	19.6000
2010.09	25.5263	25.1059	25.9466	22.8000	25.2000
2010.10	24.0791	23.6637	24.4945	21.0000	23.5000
2010.11	24.5084	24.0641	24.9528	20.9000	21.6000
2010.12	21.8645	21.4241	22.3049	13.9000	14.5000
2011.01	76.3946	74.8131	77.9761	17.7000	18.7000
2011.02	66.5058	65.0835	67.9282	29.1000	29.6000
2011.03	69.2956	67.9759	70.6154	48.0000	55.8000
2011.04	77.6410	76.2479	79.0342	47.3000	54.4000
2011.05	78.2011	76.8824	79.5197	37.3000	41.5000
2011.06	64.4044	63.2770	65.5317	35.2000	37.0000
2011.07	69.5734	68.3827	70.7640	41.5000	43.8000
2011.08	73.4346	72.2525	74.6167	42.4000	50.5000
2011.09	81.0860	79.6860	82.4860	73.8000	78.0000
2011.10	76.5486	75.2625	77.8347	78.9000	88.0000
2011.11	77.6302	76.0111	79.2493	84.6000	96.7000
2011.12	68.2225	66.8192	69.6257	65.8000	73.0000
2012.01	82.0270	80.4195	83.6344	55.8000	58.2000
2012.02	70.2080	68.7839	71.6322	29.2000	33.1000
2012.03	73.8133	72.5088	75.1178	53.1000	64.1000
2012.04	81.5007	80.0793	82.9222	51.4000	55.2000
2012.05	83.5884	82.2124	84.9644	61.8000	69.0000
2012.06	68.1872	67.0301	69.3444	59.7000	64.5000
2012.07	74.0502	72.8369	75.2634	64.2000	51.3000
2012.08	75.3679	74.1594	76.5764	57.7000	63.1000
2012.09	83.7641	82.3123	85.2159	57.7000	61.5000
2012.10	79.9382	78.4780	81.3984	48.3000	53.3000
2012.11	81.0633	79.4401	82.6865	56.7000	61.4000
2012.12	71.3986	69.8519	72.9454	37.4000	40.8000
2013.01	91.1994	89.4603	92.9384	63.8000	62.9000
2013.02	78.1942	76.6237	79.7647	37.8000	38.0000
2013.03	79.6140	78.0017	81.2264	50.6000	57.9000
2013.04	88.8812	87.3256	90.4368	70.6000	72.4000
2013.05	89.1256	87.5445	90.7066	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	74.1451	72.8377	75.4525	51.0000	52.5000
2013.07	79.5294	78.2558	80.8031	57.0000	57.0000
2013.08	82.5043	81.1814	83.8273	60.0000	66.0000
2013.09	90.4114	88.8043	92.0186	34.6000	36.9000
2013.10	85.2011	83.6271	86.7752	74.5000	85.6000
2013.11	84.8497	82.9542	86.7453	73.9000	77.6000
2013.12	76.8500	75.2153	78.4847	77.8000	90.3000
2014.01	106.2497	104.0216	108.4777	77.4000	82.0000
2014.02	92.9188	91.0939	94.7436	93.9000	102.8000
2014.03	96.8385	95.0819	98.5951	80.9000	92.2000
2014.04	108.2712	106.3884	110.1539	76.9000	84.7000
2014.05	109.2335	107.3767	111.0904	72.3000	75.2000
2014.06	90.6819	89.1530	92.2108	67.2000	71.0000
2014.07	96.9808	95.3718	98.5898	72.5000	72.5000
2014.08	100.7440	99.1748	102.3132	71.2000	74.7000
2014.09	111.5314	109.5713	113.4915	83.2000	87.6000
2014.10	104.8040	102.8793	106.7286	59.5000	60.6000
2014.11	105.3921	103.2050	107.5792	65.8000	71.1000
2014.12	93.4957	91.3367	95.6548	75.8000	78.0000
2015.01	65.7438	64.4322	67.0555	65.9000	67.0000
2015.02	56.0663	54.8344	57.2981	42.4000	44.8000
2015.03	59.1725	58.0912	60.2537	38.0000	38.4000
2015.04	65.7386	64.5600	66.9171	49.0000	54.4000
2015.05	66.6102	65.5083	67.7122	56.3000	58.8000
2015.06	55.2138	54.2176	56.2099	50.2000	68.3000
2015.07	58.5066	57.5074	59.5058	47.9000	65.8000
2015.08	62.0488	61.0053	63.0923	39.5000	57.2000
2015.09	67.7875	66.5561	69.0189	49.2000	72.1000
2015.10	64.1993	62.9599	65.4386	39.3000	48.3000
2015.11	65.2842	63.8689	66.6995	39.6000	55.9000
2015.12	57.9923	56.7201	59.2645	36.4000	44.8000
2016.01	35.9669	35.2176	36.7162	33.7000	43.3000
2016.02	30.8216	30.1803	31.4629	38.3000	46.8000
2016.03	32.0203	31.3827	32.6578	30.5000	38.9000
2016.04	35.4421	34.7668	36.1173	26.6000	30.9000
2016.05	36.0118	35.3532	36.6703	33.7000	48.4000
2016.06	29.5303	29.0265	30.0341	13.1000	19.5000
2016.07	31.8870	31.3733	32.4006	21.2000	27.5000
2016.08	33.4398	32.8510	34.0286	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	37.3948	36.7114	38.0783	27.7000	37.1000
2016.10	35.0206	34.3448	35.6963	22.7000	31.7000
2016.11	35.1987	34.4629	35.9345	14.0000	22.2000
2016.12	31.6799	31.0038	32.3559	11.1000	20.0000
2017.01	19.5791	19.1674	19.9908	18.4000	26.2000
2017.02	16.8458	16.4763	17.2154	14.4000	20.6000
2017.03	17.6528	17.3170	17.9887	11.3000	15.5000
2017.04	19.7629	19.4141	20.1117	21.6000	33.2000
2017.05	19.7558	19.4147	20.0970	12.5000	18.1000
2017.06	16.2169	15.9467	16.4871	15.5000	19.3000
2017.07	17.5853	17.3062	17.8643	11.5000	16.3000
2017.08	18.3654	18.0524	18.6784	22.8000	35.7000
2017.09	20.7533	20.3660	21.1407	34.6000	42.9000
2017.10	19.0520	18.6778	19.4262	10.5000	11.0000
2017.11	18.9907	18.5907	19.3908	4.2000	5.6000
2017.12	16.9432	16.6794	17.2070	4.0000	4.6000
2018.01	5.2242	5.1147	5.3337	3.1000	6.3000
2018.02	4.4579	4.3548	4.5609	6.8000	11.8000
2018.03	4.5995	4.5084	4.6906	1.1000	1.2000

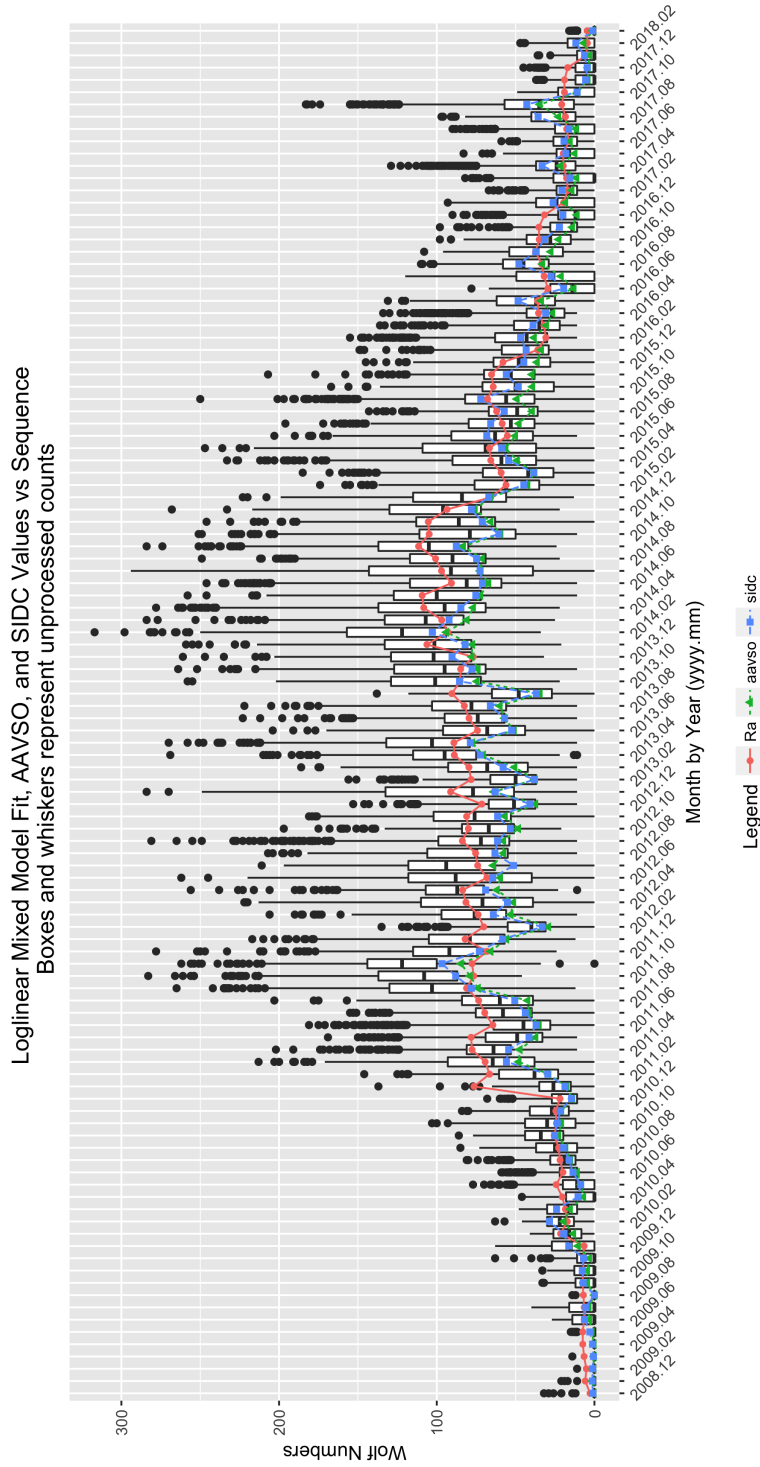


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201803 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4235	0.3195	4.4555	0.0000
seeF	-0.2173	0.0061	-35.7701	0.0000
seeG	-0.1151	0.0053	-21.7478	0.0000
seeM	-0.1926	0.0248	-7.7550	0.0000
seeP	-0.3247	0.0087	-37.3179	0.0000
sidc1	0.1205	0.0690	1.7463	0.0808
year2009	0.6449	0.3205	2.0120	0.0442
year2010	1.8557	0.3183	5.8300	0.0000
year2011	2.9767	0.3182	9.3550	0.0000
year2012	3.0143	0.3182	9.4733	0.0000
year2013	3.1102	0.3182	9.7749	0.0000
year2014	3.3073	0.3182	10.3944	0.0000
year2015	2.8227	0.3182	8.8712	0.0000
year2016	2.2056	0.3182	6.9309	0.0000
year2017	1.6016	0.3183	5.0322	0.0000
year2018	0.2705	0.3204	0.8442	0.3986
mon2	-0.1448	0.0095	-15.1681	0.0000
mon3	-0.1113	0.0090	-12.4037	0.0000
mon4	-0.0106	0.0087	-1.2252	0.2205
mon5	-0.0077	0.0085	-0.8962	0.3702
mon6	-0.2038	0.0090	-22.7265	0.0000
mon7	-0.1351	0.0087	-15.6173	0.0000
mon8	-0.0886	0.0085	-10.4279	0.0000
mon9	0.0213	0.0085	2.5001	0.0124
mon10	-0.0349	0.0087	-3.9933	0.0001
mon11	-0.0147	0.0091	-1.6084	0.1078
mon12	-0.1243	0.0093	-13.3506	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: 201803 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 1.00	Length:96858	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 :15.72		Mean :0.2638
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: 201803 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:96858	Length:96858
1st Qu.: 1.000	1st Qu.: 5.00	1st Qu.: 23.00	Class :character	Class :character
Median : 3.000	Median : 15.00	Median : 50.00	Mode :character	Mode :character
Mean : 3.657	Mean : 21.92	Mean : 58.49		
3rd Qu.: 5.000	3rd Qu.: 32.00	3rd Qu.: 87.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201803 Summary of Sunspot Numbers

inst	filter	unit
Length:96858	Length:96858	Length:96858
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201803 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 : 110.6	Mean : 16.97	Mean :1113	Mean : 187.6
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

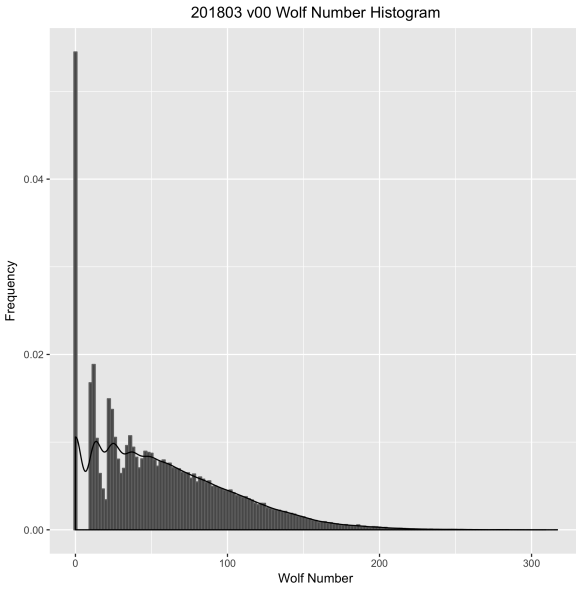


Figure 3: Box plots of raw Wolf number (w) by observer rank.

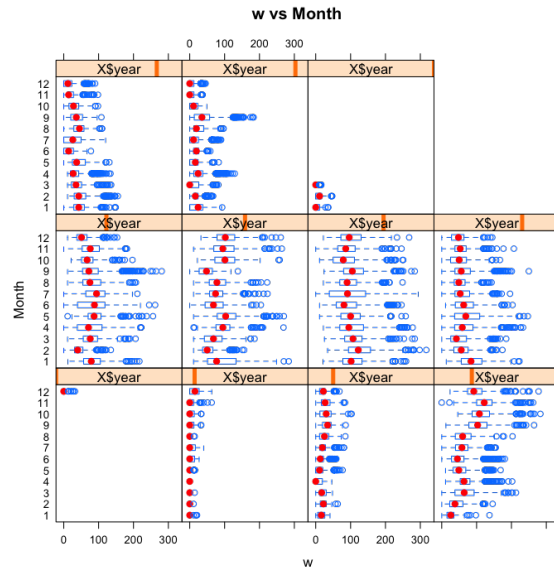


Figure 4: Box plots of raw Wolf number (w) by month and year.

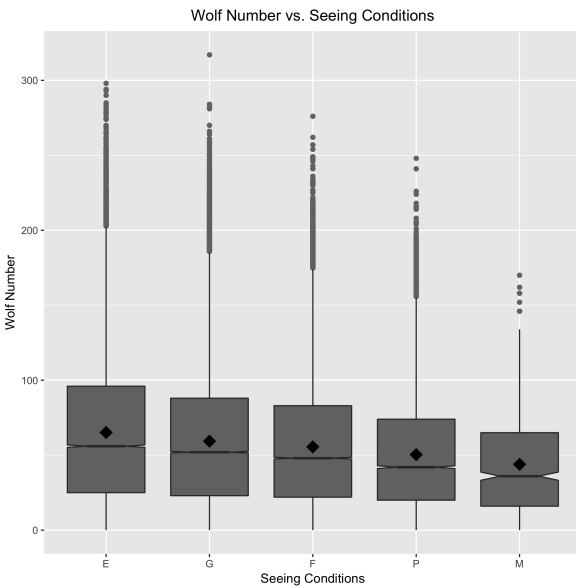


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

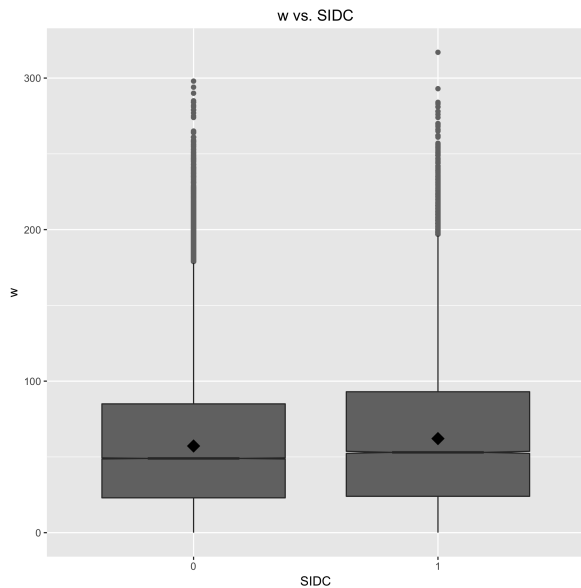


Figure 6: Box plots of raw Wolf number (w) by organization.

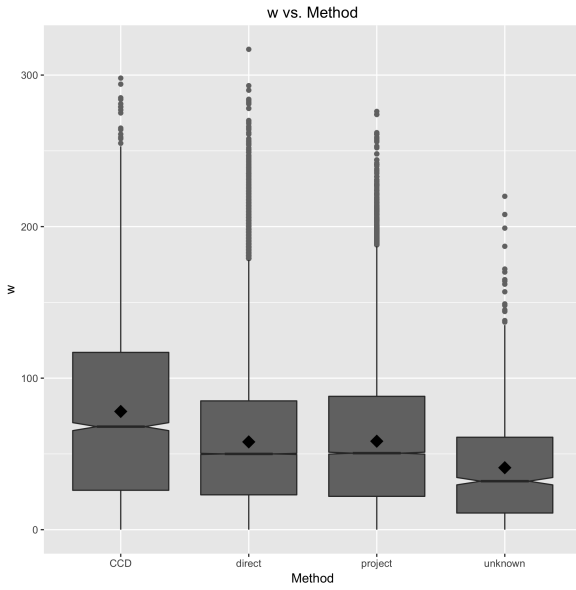


Figure 7: Box plots of raw Wolf number (w) by observer rank.

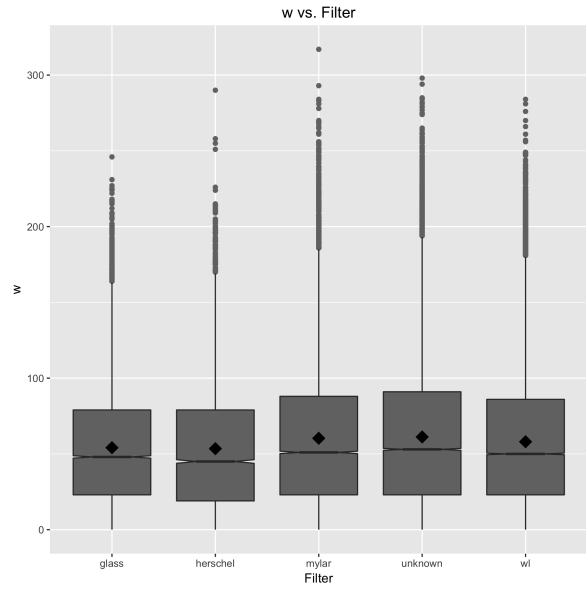


Figure 8: Box plots of raw Wolf number (w) by month and year.

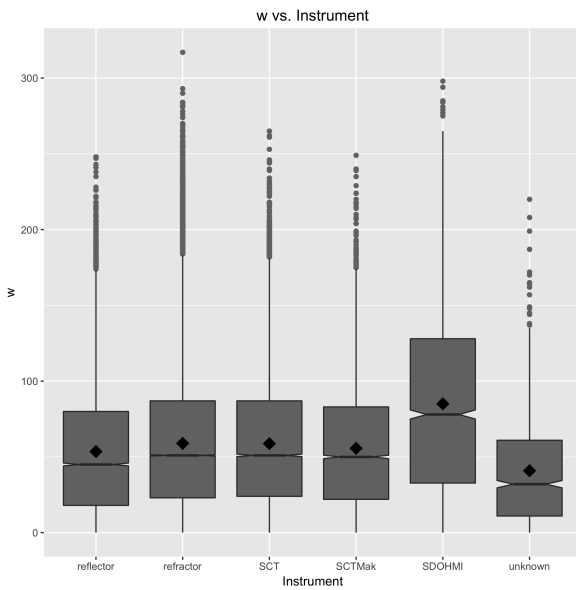


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

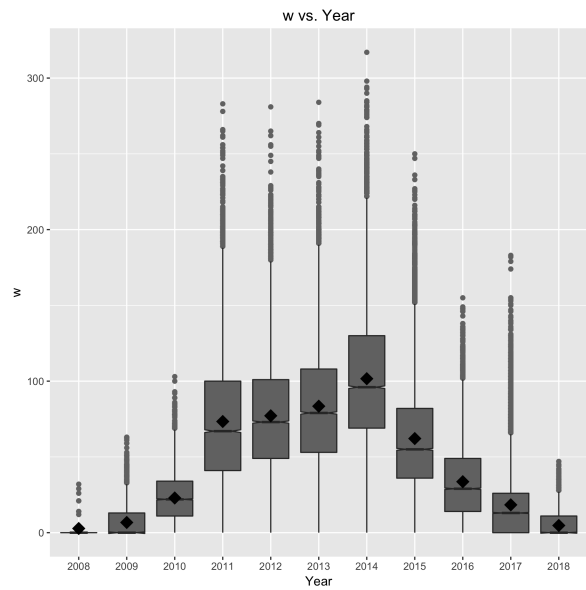


Figure 10: Box plots of raw Wolf number (w) by organization.

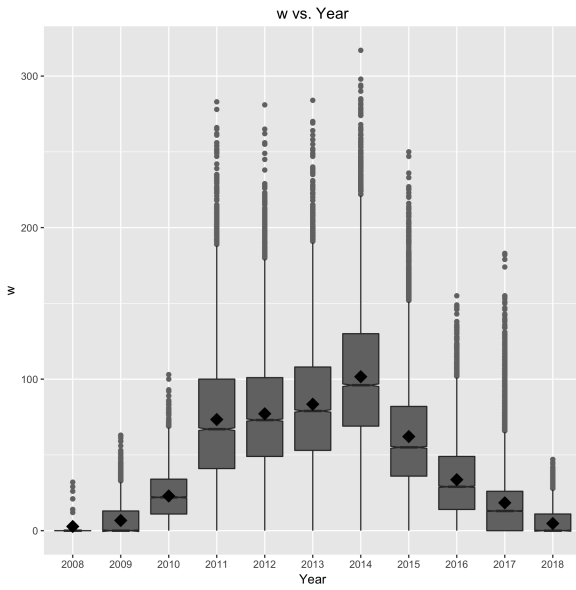


Figure 11: Box plots of raw Wolf number (w) by year.

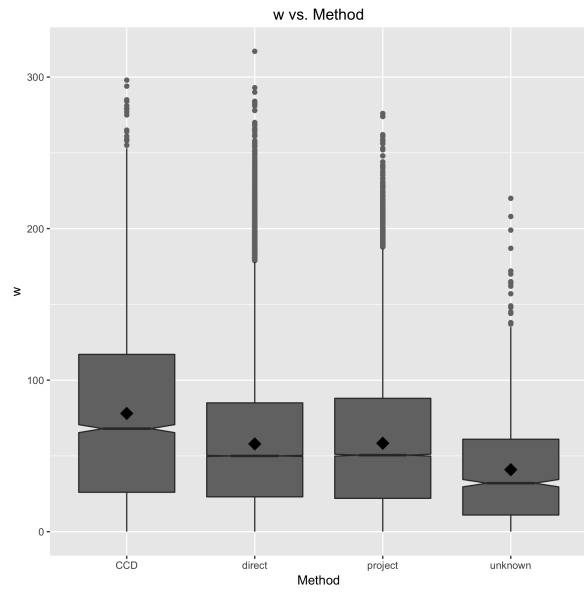


Figure 12: Box plots of raw Wolf number (w) by observing method.