Monthly Report (00)

201802 Data Set

Wednesday 14th March, 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.

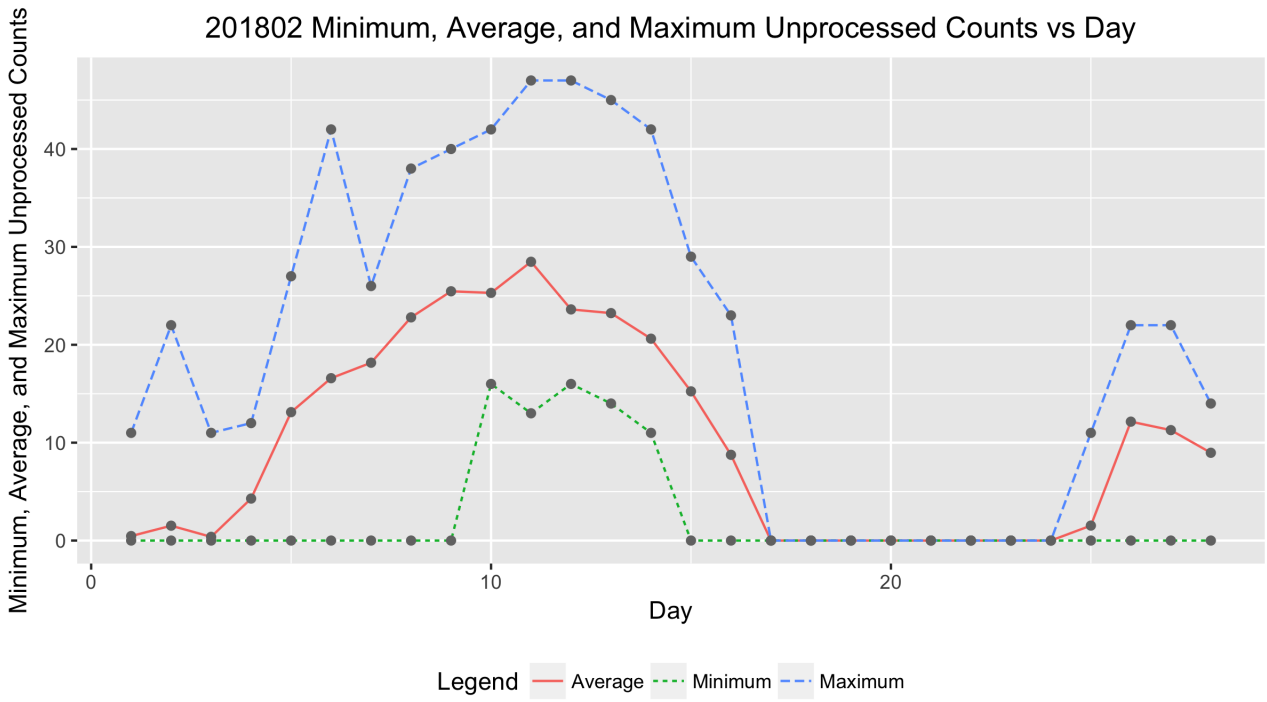

Figure 1: Raw average sunspot count by day of the month.

Table 1: 201802 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	24.0000	0.0000	0.4583	11.0000
2.0000	29.0000	0.0000	1.5172	22.0000
3.0000	29.0000	0.0000	0.3793	11.0000
4.0000	31.0000	0.0000	4.2903	12.0000
5.0000	34.0000	0.0000	13.1176	27.0000
6.0000	29.0000	0.0000	16.5862	42.0000
7.0000	30.0000	0.0000	18.1667	26.0000
8.0000	35.0000	0.0000	22.8000	38.0000
9.0000	28.0000	0.0000	25.4643	40.0000
10.0000	20.0000	16.0000	25.3000	42.0000
11.0000	25.0000	13.0000	28.4800	47.0000
12.0000	31.0000	16.0000	23.6129	47.0000
13.0000	34.0000	14.0000	23.2353	45.0000
14.0000	29.0000	11.0000	20.6207	42.0000
15.0000	29.0000	0.0000	15.2414	29.0000
16.0000	29.0000	0.0000	8.7586	23.0000
17.0000	33.0000	0.0000	0.0000	0.0000
18.0000	33.0000	0.0000	0.0000	0.0000
19.0000	20.0000	0.0000	0.0000	0.0000
20.0000	28.0000	0.0000	0.0000	0.0000
21.0000	28.0000	0.0000	0.0000	0.0000
22.0000	29.0000	0.0000	0.0000	0.0000
23.0000	31.0000	0.0000	0.0000	0.0000
24.0000	38.0000	0.0000	0.0000	0.0000
25.0000	36.0000	0.0000	1.5278	11.0000
26.0000	34.0000	0.0000	12.1471	22.0000
27.0000	39.0000	0.0000	11.2821	22.0000
28.0000	31.0000	0.0000	8.9677	14.0000

3 Error Tables

Data are for the month of February 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	2.4104	3.1305	0.5000	1.0000
2009.01	5.7879	5.1661	6.4096	1.3000	1.3000
2009.02	5.1002	4.5374	5.6630	0.7000	1.2000
2009.03	6.6293	6.3651	6.8935	0.3000	0.6000
2009.04	7.3840	7.1121	7.6559	0.4000	1.2000
2009.05	7.4192	7.1171	7.7213	1.6000	2.9000
2009.06	6.5177	6.1815	6.8538	3.2000	6.3000
2009.07	6.4896	6.2233	6.7559	3.6000	5.5000
2009.08	7.0018	6.7166	7.2870	0.0000	0.0000
2009.09	7.5775	7.3013	7.8537	4.5000	7.1000
2009.10	7.0662	6.6864	7.4459	4.5000	7.7000
2009.11	7.0253	6.8317	7.2189	3.3000	6.9000
2009.12	6.5415	6.3548	6.7282	10.4000	16.3000
2010.01	21.3653	18.9504	23.7801	13.3000	19.5000
2010.02	17.1255	14.8149	19.4360	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	18.5538	16.2678	20.8397	15.4000	24.0000
2010.04	20.4114	18.0217	22.8012	7.0000	10.4000
2010.05	24.2408	23.8126	24.6690	8.4000	8.7000
2010.06	19.9566	19.6251	20.2881	11.0000	13.6000
2010.07	21.7913	21.4760	22.1066	15.2000	16.1000
2010.08	22.7287	22.3558	23.1016	18.3000	19.6000
2010.09	25.5268	25.1065	25.9470	22.8000	25.2000
2010.10	24.0781	23.6629	24.4933	21.0000	23.5000
2010.11	24.5042	24.0600	24.9485	20.9000	21.6000
2010.12	21.8641	21.4242	22.3041	13.9000	14.5000
2011.01	76.0706	74.4976	77.6436	17.7000	18.7000
2011.02	66.2087	64.7945	67.6229	29.1000	29.6000
2011.03	69.8239	68.4946	71.1531	48.0000	55.8000
2011.04	77.6369	76.2442	79.0297	47.3000	54.4000
2011.05	78.1936	76.8752	79.5120	37.3000	41.5000
2011.06	64.4027	63.2756	65.5297	35.2000	37.0000
2011.07	69.5737	68.3834	70.7641	41.5000	43.8000
2011.08	73.4322	72.2504	74.6140	42.4000	50.5000
2011.09	81.0858	79.6861	82.4855	73.8000	78.0000
2011.10	76.5450	75.2592	77.8308	78.9000	88.0000
2011.11	77.6252	76.0066	79.2437	84.6000	96.7000
2011.12	68.2172	66.8145	69.6199	65.8000	73.0000
2012.01	81.6655	80.0658	83.2651	55.8000	58.2000
2012.02	69.8800	68.4633	71.2967	29.2000	33.1000
2012.03	74.3787	73.0648	75.6926	53.1000	64.1000
2012.04	81.4991	80.0783	82.9199	51.4000	55.2000
2012.05	83.5856	82.2102	84.9611	61.8000	69.0000
2012.06	68.1858	67.0291	69.3424	59.7000	64.5000
2012.07	74.0464	72.8337	75.2592	64.2000	51.3000
2012.08	75.3665	74.1586	76.5743	57.7000	63.1000
2012.09	83.7630	82.3119	85.2141	57.7000	61.5000
2012.10	79.9381	78.4787	81.3975	48.3000	53.3000
2012.11	81.0646	79.4423	82.6869	56.7000	61.4000
2012.12	71.3968	69.8509	72.9426	37.4000	40.8000
2013.01	90.8101	89.0796	92.5405	63.8000	62.9000
2013.02	77.8366	76.2743	79.3989	37.8000	38.0000
2013.03	80.2310	78.6071	81.8550	50.6000	57.9000
2013.04	88.8870	87.3324	90.4416	70.6000	72.4000
2013.05	89.1335	87.5534	90.7136	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	74.1471	72.8406	75.4537	51.0000	52.5000
2013.07	79.5302	78.2574	80.8030	57.0000	57.0000
2013.08	82.5064	81.1841	83.8288	60.0000	66.0000
2013.09	90.4181	88.8118	92.0244	34.6000	36.9000
2013.10	85.2047	83.6315	86.7779	74.5000	85.6000
2013.11	84.8550	82.9606	86.7495	73.9000	77.6000
2013.12	76.8513	75.2175	78.4851	77.8000	90.3000
2014.01	105.7948	103.5779	108.0116	77.4000	82.0000
2014.02	92.4914	90.6759	94.3068	93.9000	102.8000
2014.03	97.5835	95.8143	99.3527	80.9000	92.2000
2014.04	108.2702	106.3887	110.1517	76.9000	84.7000
2014.05	109.2330	107.3771	111.0889	72.3000	75.2000
2014.06	90.6814	89.1532	92.2095	67.2000	71.0000
2014.07	96.9825	95.3743	98.5907	72.5000	72.5000
2014.08	100.7453	99.1769	102.3136	71.2000	74.7000
2014.09	111.5343	109.5750	113.4936	83.2000	87.6000
2014.10	104.8034	102.8795	106.7273	59.5000	60.6000
2014.11	105.3945	103.2089	107.5801	65.8000	71.1000
2014.12	93.4932	91.3356	95.6509	75.8000	78.0000
2015.01	65.4602	64.1550	66.7655	65.9000	67.0000
2015.02	55.8061	54.5808	57.0315	42.4000	44.8000
2015.03	59.6244	58.5356	60.7133	38.0000	38.4000
2015.04	65.7380	64.5600	66.9160	49.0000	54.4000
2015.05	66.6090	65.5075	67.7104	56.3000	58.8000
2015.06	55.2135	54.2178	56.2091	50.2000	68.3000
2015.07	58.5060	57.5074	59.5047	47.9000	65.8000
2015.08	62.0496	61.0067	63.0926	39.5000	57.2000
2015.09	67.7887	66.5579	69.0195	49.2000	72.1000
2015.10	64.1974	62.9587	65.4362	39.3000	48.3000
2015.11	65.2793	63.8650	66.6936	39.6000	55.9000
2015.12	57.9873	56.7160	59.2586	36.4000	44.8000
2016.01	35.8124	35.0668	36.5580	33.7000	43.3000
2016.02	30.6801	30.0421	31.3181	38.3000	46.8000
2016.03	32.2689	31.6269	32.9109	30.5000	38.9000
2016.04	35.4412	34.7665	36.1160	26.6000	30.9000
2016.05	36.0115	35.3533	36.6697	33.7000	48.4000
2016.06	29.5315	29.0280	30.0351	13.1000	19.5000
2016.07	31.8882	31.3747	32.4016	21.2000	27.5000
2016.08	33.4408	32.8522	34.0293	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	37.3961	36.7130	38.0792	27.7000	37.1000
2016.10	35.0233	34.3479	35.6987	22.7000	31.7000
2016.11	35.2006	34.4652	35.9360	14.0000	22.2000
2016.12	31.6819	31.0061	32.3577	11.1000	20.0000
2017.01	19.4948	19.0850	19.9045	18.4000	26.2000
2017.02	16.7689	16.4012	17.1366	14.4000	20.6000
2017.03	17.7902	17.4519	18.1284	11.3000	15.5000
2017.04	19.7637	19.4151	20.1123	21.6000	33.2000
2017.05	19.7579	19.4169	20.0989	12.5000	18.1000
2017.06	16.2172	15.9471	16.4872	15.5000	19.3000
2017.07	17.5852	17.3063	17.8640	11.5000	16.3000
2017.08	18.3631	18.0506	18.6756	22.8000	35.7000
2017.09	20.7422	20.3581	21.1263	34.6000	42.9000
2017.10	19.0524	18.6794	19.4254	10.5000	11.0000
2017.11	18.9916	18.5922	19.3910	4.2000	5.6000
2017.12	16.9482	16.6852	17.2111	4.0000	4.6000
2018.01	7.3594	7.2066	7.5123	3.1000	6.3000
2018.02	6.2951	6.1512	6.4390	6.8000	11.8000

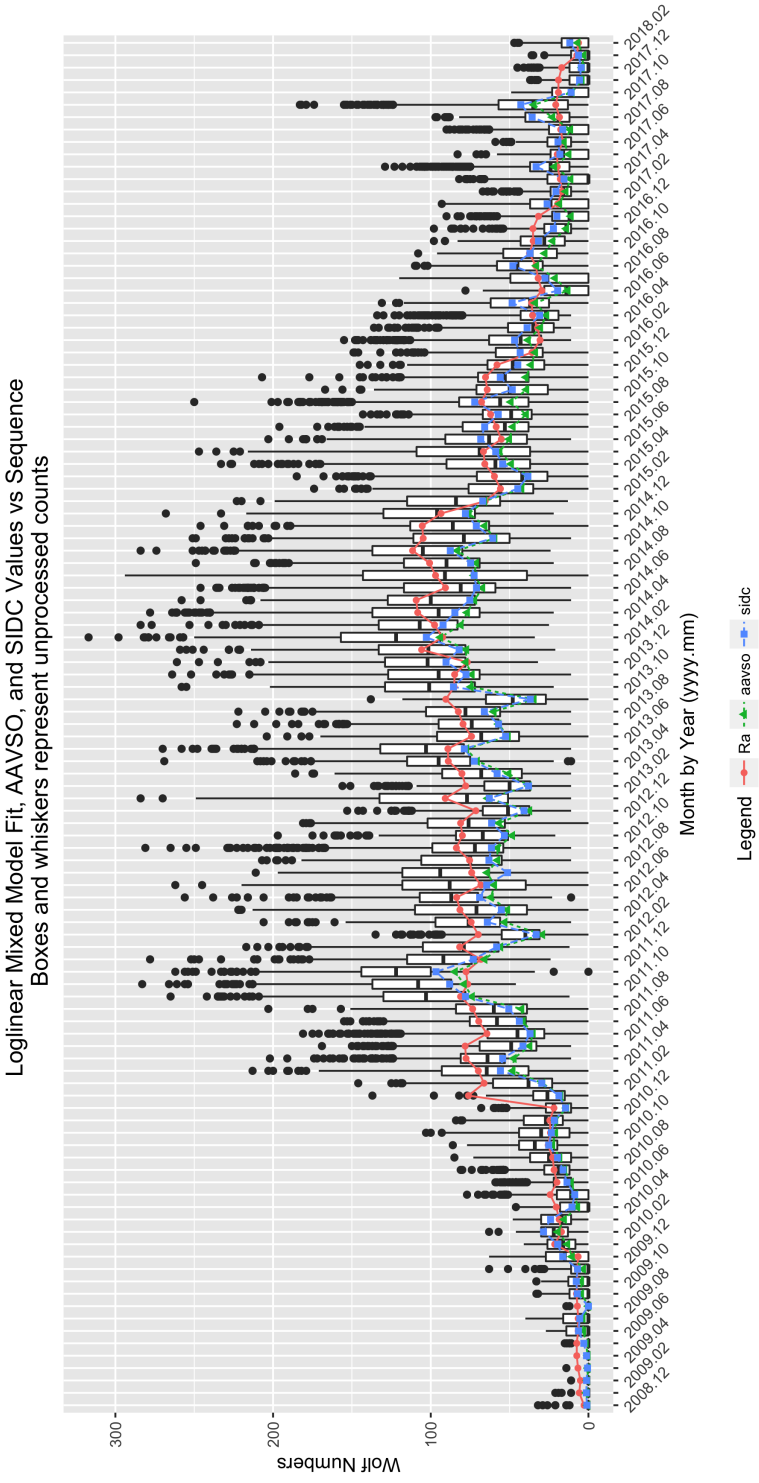

Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201802 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	1.4223	0.3202	4.4426	0.0000
seeF	-0.2173	0.0061	-35.6945	0.0000
seeG	-0.1152	0.0053	-21.7317	0.0000
seeM	-0.1927	0.0249	-7.7437	0.0000
seeP	-0.3245	0.0087	-37.2263	0.0000
sidc1	0.1113	0.0685	1.6250	0.1042
year2009	0.6451	0.3212	2.0085	0.0446
year2010	1.8558	0.3189	5.8186	0.0000
year2011	2.9769	0.3188	9.3365	0.0000
year2012	3.0145	0.3188	9.4546	0.0000
year2013	3.1105	0.3188	9.7557	0.0000
year2014	3.3076	0.3188	10.3739	0.0000
year2015	2.8229	0.3188	8.8535	0.0000
year2016	2.2059	0.3189	6.9175	0.0000
year2017	1.6018	0.3189	5.0227	0.0000
year2018	0.6167	0.3213	1.9196	0.0549
mon2	-0.1450	0.0096	-15.1656	0.0000
mon3	-0.0992	0.0090	-11.0072	0.0000
mon4	-0.0063	0.0087	-0.7205	0.4712
mon5	-0.0033	0.0086	-0.3804	0.7036
mon6	-0.1994	0.0090	-22.1679	0.0000
mon7	-0.1307	0.0087	-15.0629	0.0000
mon8	-0.0842	0.0085	-9.8797	0.0000
mon9	0.0257	0.0085	3.0081	0.0026
mon10	-0.0306	0.0088	-3.4830	0.0005
mon11	-0.0104	0.0092	-1.1309	0.2581
mon12	-0.1200	0.0093	-12.8465	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: 201802 Summary of Sunspot Numbers

jd	year	mon	day	obs
Min. :2454802	Min. :2008	Min. : 1.000	Min. : 1.00	Length:95893
1st Qu.:2456080	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00	Class :character
Median :2456772	Median :2014	Median : 7.000	Median :16.00	Mode :character
Mean :2456760	Mean :2014	Mean : 6.648	Mean :15.71	
3rd Qu.:2457483	3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.00	
Max. :2458178	Max. :2018	Max. :12.000	Max. :31.00	

Table 5: Summary of Sunspot Numbers

g	s	w	aperture
Min. : 0.000	Min. : 0.00	Min. : 0.00	Min. : 0.0
1st Qu.: 2.000	1st Qu.: 5.00	1st Qu.: 23.00	1st Qu.: 76.0
Median : 3.000	Median : 15.00	Median : 51.00	Median : 90.0
Mean : 3.692	Mean : 22.14	Mean : 59.07	Mean : 110.7
3rd Qu.: 5.000	3rd Qu.: 32.00	3rd Qu.: 87.00	3rd Qu.: 125.0
Max. :19.000	Max. :204.00	Max. :317.00	Max. :1524.0

Table 6: Summary of Sunspot Numbers

eyep	foclen	mag	k	ow
Min. : 0.00	Min. : 0	Min. : 0.0	Min. :0.0000	Min. :0.0000
1st Qu.: 2.00	1st Qu.: 800	1st Qu.: 40.0	1st Qu.:0.0000	1st Qu.:0.0000
Median : 13.00	Median :1000	Median : 57.5	Median :0.5810	Median :1.0000
Mean : 16.36	Mean :1115	Mean : 187.5	Mean :0.4086	Mean :0.5726
3rd Qu.: 23.00	3rd Qu.:1296	3rd Qu.: 77.0	3rd Qu.:0.6910	3rd Qu.:1.0000
Max. :2010.00	Max. :4300	Max. :4591.0	Max. :1.2160	Max. :1.0000

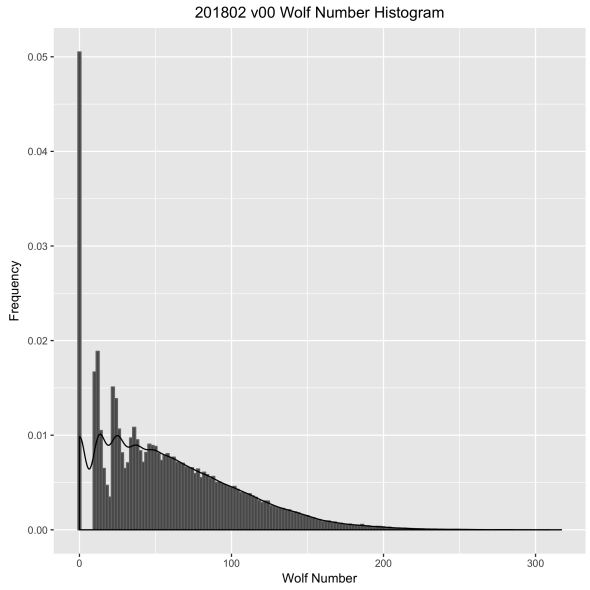

Figure 3: Box plots of raw Wolf number (w) by observer rank.

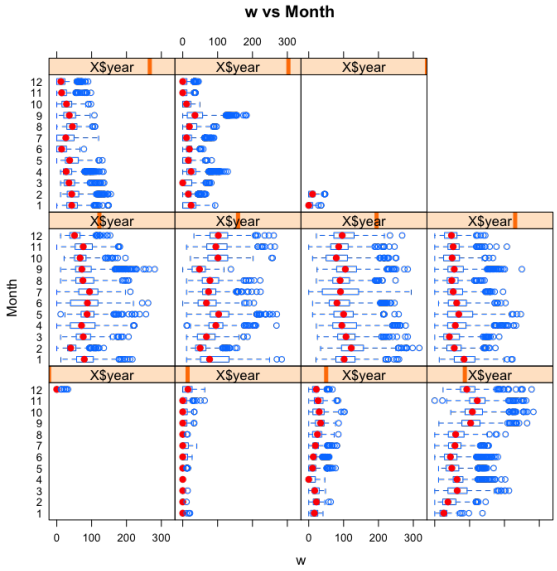

Figure 4: Box plots of raw Wolf number (w) by month and year.

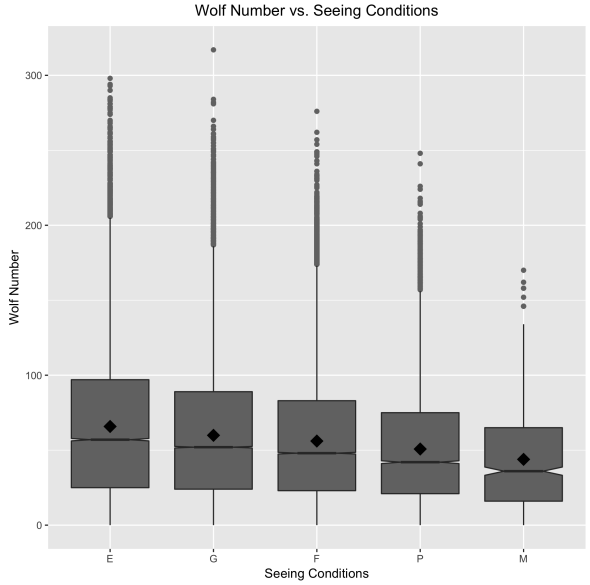

Figure 5: Box plots of raw Wolf number (w) by seeing condition.

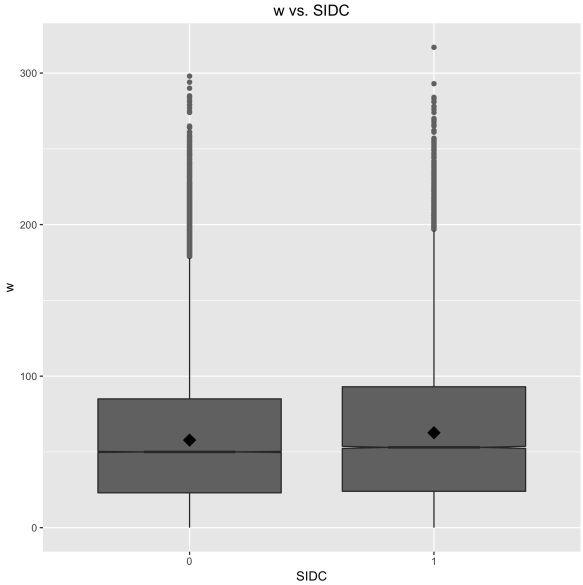

Figure 6: Box plots of raw Wolf number (w) by organization.

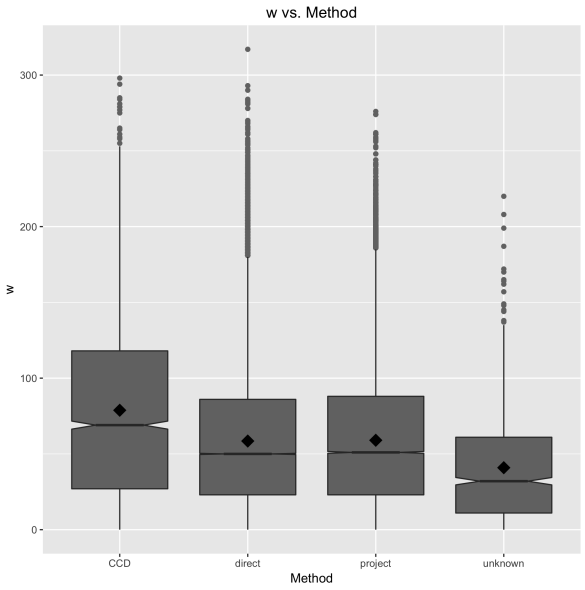

Figure 7: Box plots of raw Wolf number (w) by observer rank.

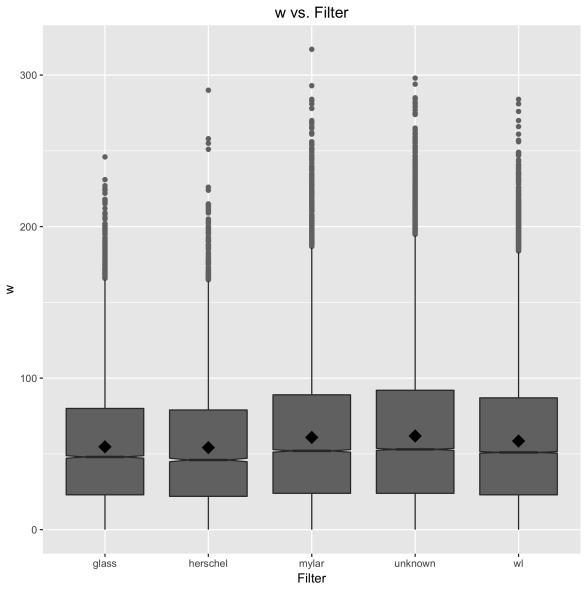

Figure 8: Box plots of raw Wolf number (w) by month and year.

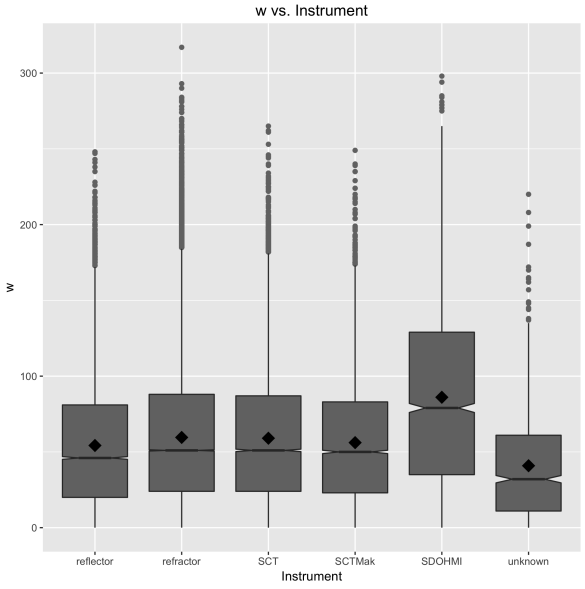

Figure 9: Box plots of raw Wolf number (w) by seeing condition.

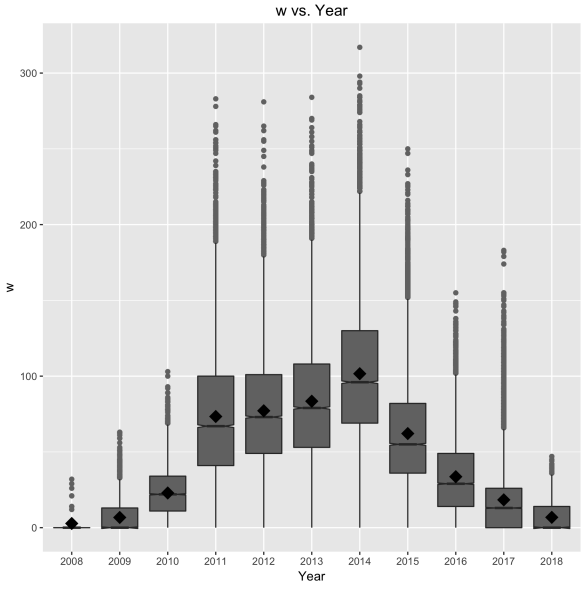

Figure 10: Box plots of raw Wolf number (w) by organization.

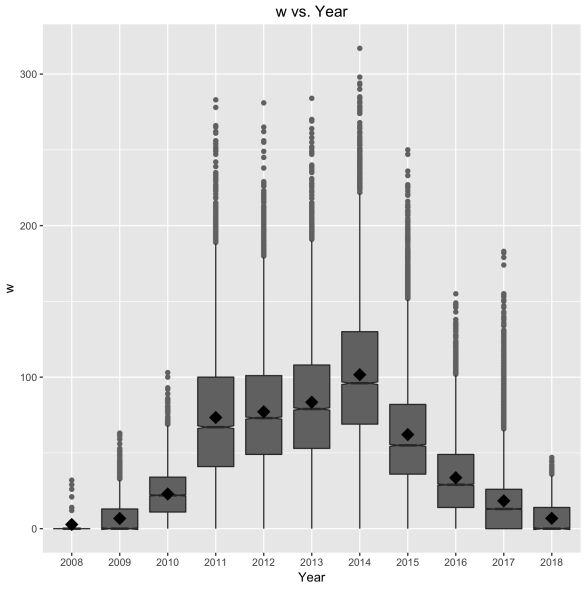

Figure 11: Box plots of raw Wolf number (w) by year.

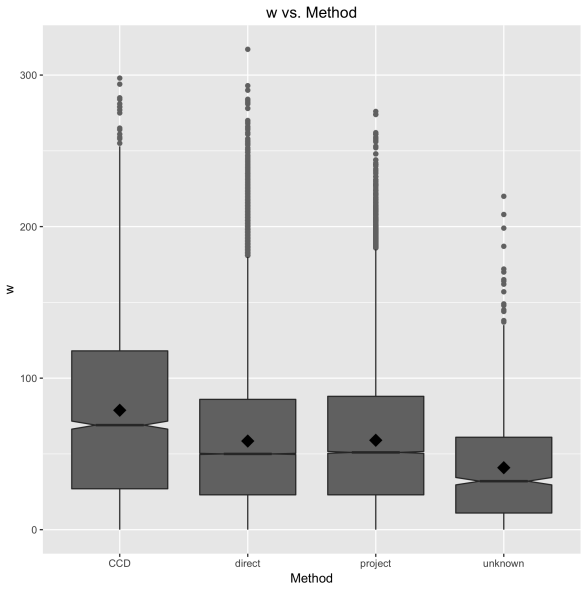

Figure 12: Box plots of raw Wolf number (w) by observing method.