

Monthly Report (00)
201902 Data Set

Tuesday 12th March, 2019

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.

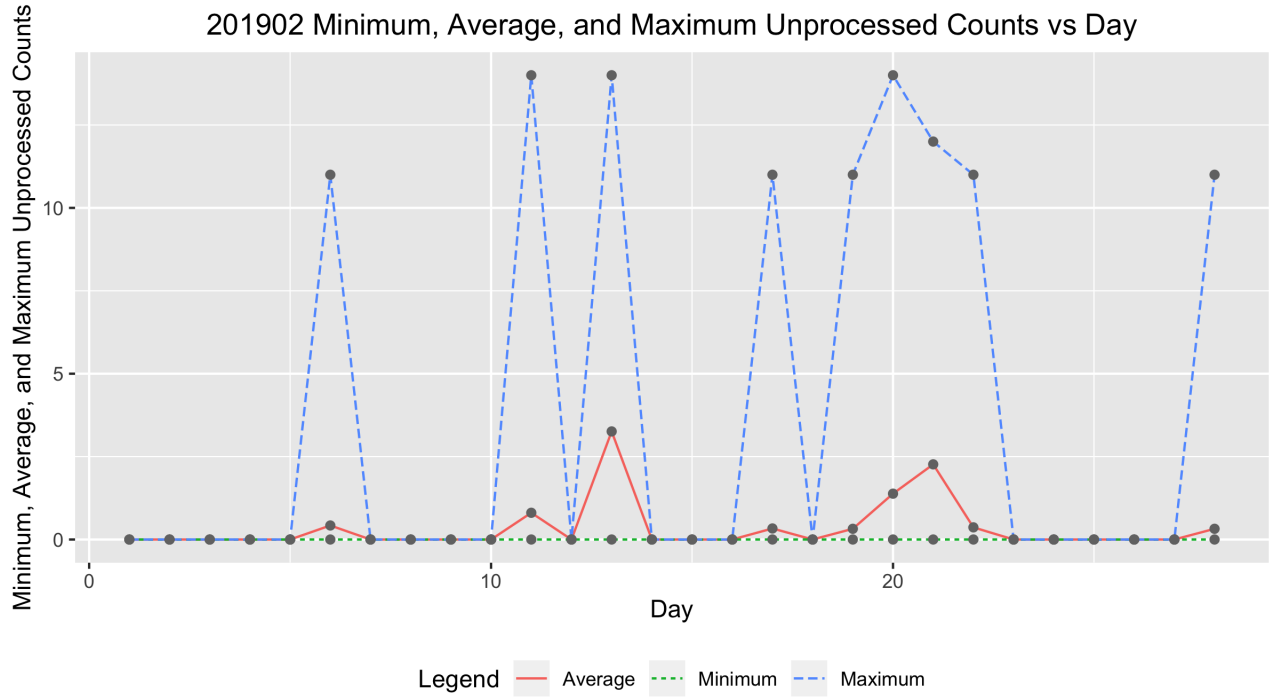


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201902 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	30.0000	0.0000	0.0000	0.0000
2.0000	27.0000	0.0000	0.0000	0.0000
3.0000	29.0000	0.0000	0.0000	0.0000
4.0000	25.0000	0.0000	0.0000	0.0000
5.0000	29.0000	0.0000	0.0000	0.0000
6.0000	26.0000	0.0000	0.4231	11.0000
7.0000	29.0000	0.0000	0.0000	0.0000
8.0000	30.0000	0.0000	0.0000	0.0000
9.0000	38.0000	0.0000	0.0000	0.0000
10.0000	26.0000	0.0000	0.0000	0.0000
11.0000	31.0000	0.0000	0.8065	14.0000
12.0000	29.0000	0.0000	0.0000	0.0000
13.0000	27.0000	0.0000	3.2593	14.0000
14.0000	33.0000	0.0000	0.0000	0.0000
15.0000	34.0000	0.0000	0.0000	0.0000
16.0000	40.0000	0.0000	0.0000	0.0000
17.0000	33.0000	0.0000	0.3333	11.0000
18.0000	32.0000	0.0000	0.0000	0.0000
19.0000	34.0000	0.0000	0.3235	11.0000
20.0000	34.0000	0.0000	1.3824	14.0000
21.0000	30.0000	0.0000	2.2667	12.0000
22.0000	30.0000	0.0000	0.3667	11.0000
23.0000	36.0000	0.0000	0.0000	0.0000
24.0000	40.0000	0.0000	0.0000	0.0000
25.0000	38.0000	0.0000	0.0000	0.0000
26.0000	30.0000	0.0000	0.0000	0.0000
27.0000	26.0000	0.0000	0.0000	0.0000
28.0000	34.0000	0.0000	0.3235	11.0000

3 Error Tables

Data are for the month of February 2019. 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	2.4075	3.1335	0.5000	1.0000
2009.01	5.8620	5.2268	6.4973	1.3000	1.3000
2009.02	5.0828	4.5170	5.6487	0.7000	1.2000
2009.03	6.5960	6.3334	6.8586	0.3000	0.6000
2009.04	7.4184	7.1457	7.6911	0.4000	1.2000
2009.05	7.5248	7.2187	7.8308	1.6000	2.9000
2009.06	6.6882	6.3432	7.0333	3.2000	6.3000
2009.07	6.3807	6.1189	6.6425	3.6000	5.5000
2009.08	7.0634	6.7762	7.3505	0.0000	0.0000
2009.09	7.5518	7.2769	7.8266	4.5000	7.1000
2009.10	7.0568	6.6778	7.4358	4.5000	7.7000
2009.11	7.0183	6.8243	7.2123	3.3000	6.9000
2009.12	6.5101	6.3240	6.6963	10.4000	16.3000
2010.01	21.6311	19.1665	24.0956	13.3000	19.5000
2010.02	17.0405	14.7212	19.3599	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	18.4123	16.1248	20.6998	15.4000	24.0000
2010.04	20.4492	18.0352	22.8632	7.0000	10.4000
2010.05	24.5101	24.0747	24.9454	8.4000	8.7000
2010.06	20.4252	20.0849	20.7655	11.0000	13.6000
2010.07	21.3652	21.0548	21.6756	15.2000	16.1000
2010.08	22.8598	22.4840	23.2356	18.3000	19.6000
2010.09	25.3780	24.9589	25.7971	22.8000	25.2000
2010.10	23.9933	23.5783	24.4084	21.0000	23.5000
2010.11	24.4468	24.0022	24.8914	20.9000	21.6000
2010.12	21.7428	21.3035	22.1821	13.9000	14.5000
2011.01	76.9540	75.3574	78.5505	17.7000	18.7000
2011.02	65.8794	64.4691	67.2897	29.1000	29.6000
2011.03	69.2748	67.9506	70.5990	48.0000	55.8000
2011.04	77.7773	76.3786	79.1760	47.3000	54.4000
2011.05	79.0464	77.7080	80.3847	37.3000	41.5000
2011.06	65.8643	64.7081	67.0206	35.2000	37.0000
2011.07	68.1848	67.0154	69.3542	41.5000	43.8000
2011.08	73.8218	72.6305	75.0132	42.4000	50.5000
2011.09	80.5711	79.1759	81.9663	73.8000	78.0000
2011.10	76.2133	74.9286	77.4979	78.9000	88.0000
2011.11	77.3996	75.7803	79.0188	84.6000	96.7000
2011.12	67.7754	66.3773	69.1736	65.8000	73.0000
2012.01	82.6110	80.9853	84.2367	55.8000	58.2000
2012.02	69.5287	68.1136	70.9438	29.2000	33.1000
2012.03	73.7741	72.4661	75.0821	53.1000	64.1000
2012.04	81.6088	80.1808	83.0367	51.4000	55.2000
2012.05	84.5162	83.1213	85.9112	61.8000	69.0000
2012.06	69.7317	68.5450	70.9183	59.7000	64.5000
2012.07	72.5826	71.3906	73.7746	64.2000	51.3000
2012.08	75.7818	74.5639	76.9996	57.7000	63.1000
2012.09	83.2428	81.7970	84.6887	57.7000	61.5000
2012.10	79.5904	78.1329	81.0479	48.3000	53.3000
2012.11	80.8186	79.1955	82.4416	56.7000	61.4000
2012.12	70.9344	69.3935	72.4753	37.4000	40.8000
2013.01	91.8502	90.0934	93.6071	63.8000	62.9000
2013.02	77.4662	75.9063	79.0260	37.8000	38.0000
2013.03	79.5927	77.9748	81.2106	50.6000	57.9000
2013.04	89.0026	87.4382	90.5670	70.6000	72.4000
2013.05	90.0887	88.4853	91.6920	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	75.8433	74.5017	77.1850	51.0000	52.5000
2013.07	77.9621	76.7099	79.2143	57.0000	57.0000
2013.08	82.9577	81.6238	84.2916	60.0000	66.0000
2013.09	89.8338	88.2310	91.4366	34.6000	36.9000
2013.10	84.8358	83.2626	86.4090	74.5000	85.6000
2013.11	84.5792	82.6819	86.4765	73.9000	77.6000
2013.12	76.3469	74.7162	77.9775	77.8000	90.3000
2014.01	106.9979	104.7453	109.2505	77.4000	82.0000
2014.02	92.0423	90.2286	93.8559	93.9000	102.8000
2014.03	96.7993	95.0363	98.5624	80.9000	92.2000
2014.04	108.4087	106.5155	110.3019	76.9000	84.7000
2014.05	110.4093	108.5245	112.2941	72.3000	75.2000
2014.06	92.7405	91.1715	94.3095	67.2000	71.0000
2014.07	95.0400	93.4576	96.6224	72.5000	72.5000
2014.08	101.2916	99.7081	102.8750	71.2000	74.7000
2014.09	110.8291	108.8744	112.7837	83.2000	87.6000
2014.10	104.3470	102.4245	106.2695	59.5000	60.6000
2014.11	105.0632	102.8734	107.2531	65.8000	71.1000
2014.12	92.8958	90.7403	95.0513	75.8000	78.0000
2015.01	66.1844	64.8576	67.5112	65.9000	67.0000
2015.02	55.5218	54.2966	56.7469	42.4000	44.8000
2015.03	59.1601	58.0750	60.2452	38.0000	38.4000
2015.04	65.8155	64.6325	66.9984	49.0000	54.4000
2015.05	67.3411	66.2239	68.4583	56.3000	58.8000
2015.06	56.4625	55.4408	57.4842	50.2000	68.3000
2015.07	57.3586	56.3762	58.3409	47.9000	65.8000
2015.08	62.3800	61.3283	63.4316	39.5000	57.2000
2015.09	67.3627	66.1349	68.5906	49.2000	72.1000
2015.10	63.9238	62.6857	65.1619	39.3000	48.3000
2015.11	65.0839	63.6670	66.5007	39.6000	55.9000
2015.12	57.6202	56.3518	58.8885	36.4000	44.8000
2016.01	36.2245	35.4667	36.9823	33.7000	43.3000
2016.02	30.5336	29.8957	31.1715	38.3000	46.8000
2016.03	32.0102	31.3699	32.6505	30.5000	38.9000
2016.04	35.5098	34.8309	36.1887	26.6000	30.9000
2016.05	36.4213	35.7534	37.0893	33.7000	48.4000
2016.06	30.2043	29.6873	30.7213	13.1000	19.5000
2016.07	31.2517	30.7475	31.7560	21.2000	27.5000
2016.08	33.6182	33.0248	34.2116	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	37.1658	36.4850	37.8466	27.7000	37.1000
2016.10	34.8659	34.1910	35.5407	22.7000	31.7000
2016.11	35.1001	34.3641	35.8361	14.0000	22.2000
2016.12	31.4687	30.7951	32.1424	11.1000	20.0000
2017.01	19.7111	19.2956	20.1266	18.4000	26.2000
2017.02	16.6807	16.3136	17.0477	14.4000	20.6000
2017.03	17.6405	17.3038	17.9773	11.3000	15.5000
2017.04	19.7859	19.4359	20.1358	21.6000	33.2000
2017.05	19.9648	19.6191	20.3104	12.5000	18.1000
2017.06	16.5711	16.2937	16.8485	15.5000	19.3000
2017.07	17.2209	16.9461	17.4956	11.5000	16.3000
2017.08	18.4598	18.1407	18.7788	22.8000	35.7000
2017.09	20.6839	20.2742	21.0935	34.6000	42.9000
2017.10	18.9710	18.5908	19.3511	10.5000	11.0000
2017.11	18.9514	18.5542	19.3485	4.2000	5.6000
2017.12	16.8893	16.6317	17.1468	4.0000	4.6000
2018.01	5.4641	5.3457	5.5824	3.1000	6.3000
2018.02	4.5822	4.4703	4.6940	6.8000	11.8000
2018.03	4.7617	4.6655	4.8579	1.1000	1.2000
2018.04	5.2931	5.1861	5.4000	4.7000	7.5000
2018.05	5.4647	5.3579	5.5714	8.4000	14.0000
2018.06	4.5264	4.4431	4.6098	10.2000	13.6000
2018.07	4.6845	4.6298	4.7393	0.5000	1.7000
2018.08	4.9705	4.8798	5.0612	5.9000	9.5000
2018.09	5.3684	5.2643	5.4725	1.6000	2.9000
2018.10	5.1926	5.0895	5.2957	2.5000	5.6000
2018.11	5.2038	5.0924	5.3153	3.1000	4.2000
2018.12	4.7531	4.6546	4.8517	1.6000	2.3000
2019.01	4.4811	4.3930	4.5692	5.4000	2.3000
2019.02	3.8527	3.7722	3.9333	0.1000	1.2000

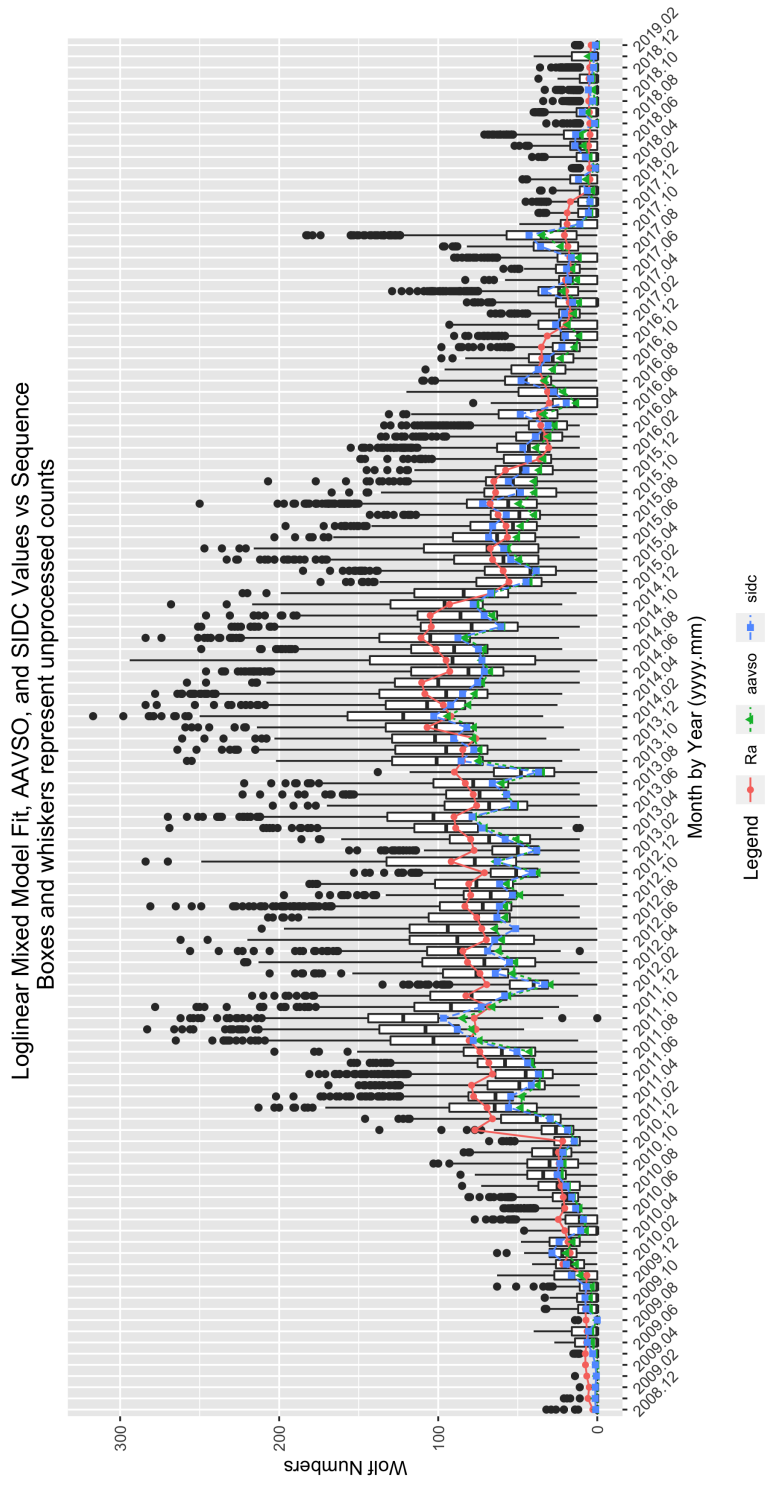


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201902 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4394	0.3152	4.5672	0.0000
seeF	-0.2186	0.0060	-36.6835	0.0000
seeG	-0.1160	0.0052	-22.3563	0.0000
seeM	-0.1975	0.0245	-8.0689	0.0000
seeP	-0.3249	0.0085	-38.0615	0.0000
sidc1	0.1636	0.0698	2.3435	0.0191
year2009	0.6399	0.3161	2.0240	0.0430
year2010	1.8494	0.3139	5.8911	0.0000
year2011	2.9698	0.3138	9.4630	0.0000
year2012	3.0073	0.3138	9.5824	0.0000
year2013	3.1033	0.3138	9.8884	0.0000
year2014	3.3003	0.3138	10.5163	0.0000
year2015	2.8157	0.3138	8.9718	0.0000
year2016	2.1988	0.3139	7.0053	0.0000
year2017	1.5940	0.3139	5.0780	0.0000
year2018	0.3019	0.3142	0.9608	0.3366
year2019	0.0962	0.3173	0.3031	0.7618
mon2	-0.1615	0.0094	-17.2286	0.0000
mon3	-0.1188	0.0088	-13.4628	0.0000
mon4	-0.0163	0.0085	-1.9174	0.0552
mon5	-0.0039	0.0083	-0.4640	0.6427
mon6	-0.1885	0.0087	-21.5688	0.0000
mon7	-0.1624	0.0085	-19.1248	0.0000
mon8	-0.0904	0.0083	-10.8785	0.0000
mon9	0.0076	0.0083	0.9115	0.3620
mon10	-0.0465	0.0086	-5.4190	0.0000
mon11	-0.0248	0.0090	-2.7681	0.0056
mon12	-0.1380	0.0091	-15.0985	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: 201902 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 0.00	Length:109669	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 : 6.635	Mean :15.73		Mean :0.2617
3rd Qu.:2017	3rd Qu.: 9.000	3rd Qu.:23.00		3rd Qu.:1.0000
Max. :2019	Max. :12.000	Max. :31.00		Max. :1.0000

Table 5: 201902 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:109669	Length:109669
1st Qu.: 1.000	1st Qu.: 2.00	1st Qu.: 13.00	Class :character	Class :character
Median : 3.000	Median : 12.00	Median : 42.00	Mode :character	Mode :character
Mean : 3.268	Mean : 19.55	Mean : 52.23		
3rd Qu.: 5.000	3rd Qu.: 29.00	3rd Qu.: 81.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201902 Summary of Sunspot Numbers

inst	filter	unit
Length:109669	Length:109669	Length:109669
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201902 Summary of Sunspot Numbers

aperture	eyep	foclen	mag
Min. : 0	Min. : 0.00	Min. : 0	Min. : 0.0
1st Qu.: 75	1st Qu.: 3.00	1st Qu.: 630	1st Qu.: 40.0
Median : 89	Median : 13.00	Median : 910	Median : 57.5
Mean : 104	Mean : 23.28	Mean :1038	Mean : 186.7
3rd Qu.: 120	3rd Qu.: 23.00	3rd Qu.:1250	3rd Qu.: 76.0
Max. :1524	Max. :2010.00	Max. :4300	Max. :4591.0

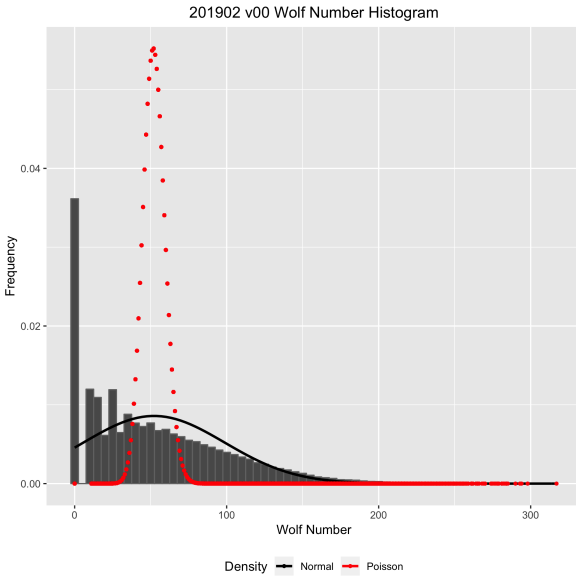


Figure 3: Box plots of raw Wolf number (w) by observer rank.

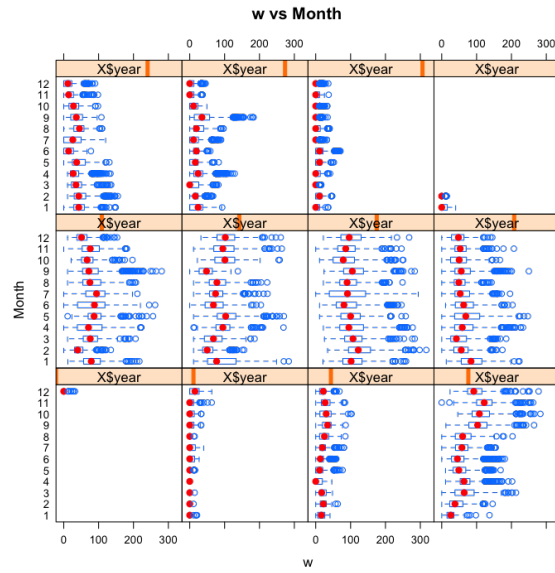


Figure 4: Box plots of raw Wolf number (w) by month and year.

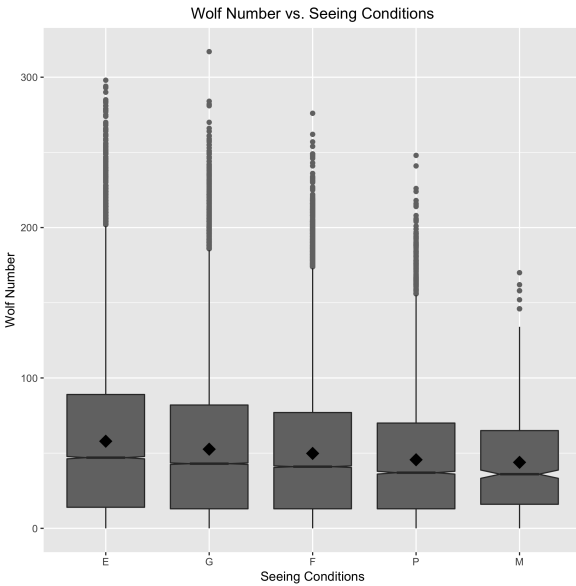


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

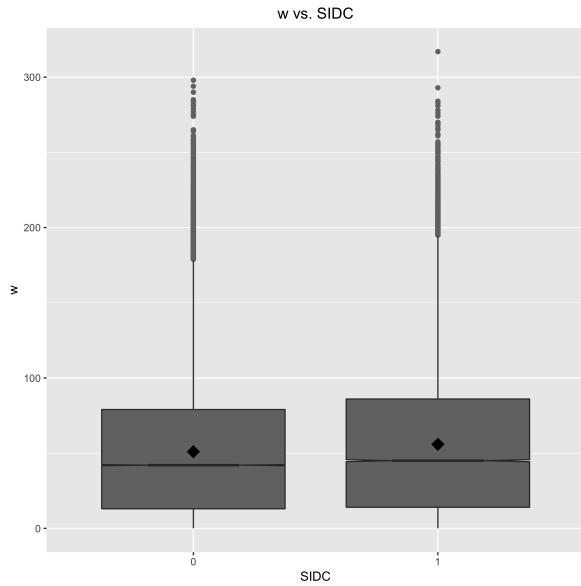


Figure 6: Box plots of raw Wolf number (w) by organization.

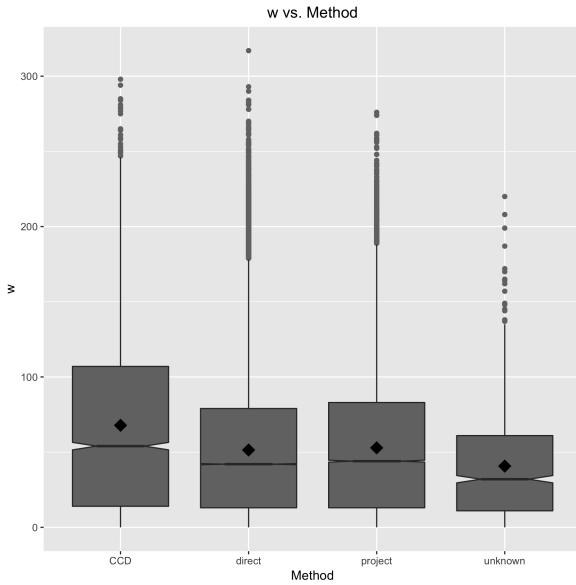


Figure 7: Box plots of raw Wolf number (w) by observer rank.

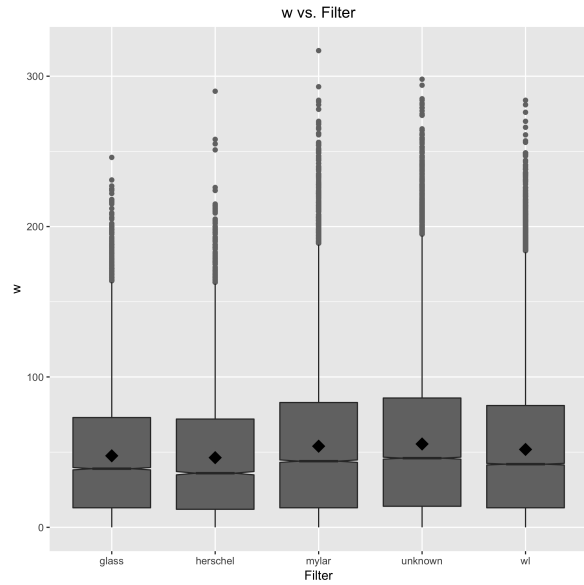


Figure 8: Box plots of raw Wolf number (w) by month and year.

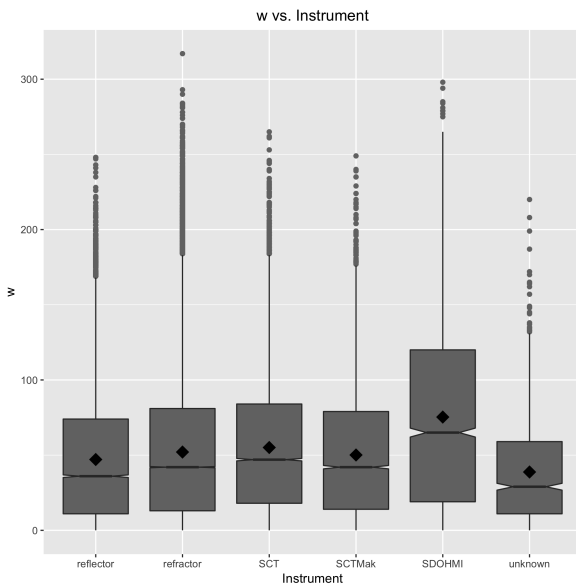


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

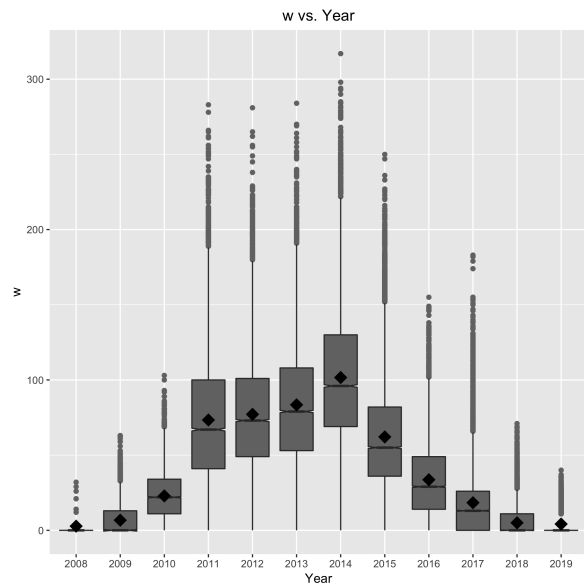


Figure 10: Box plots of raw Wolf number (w) by organization.

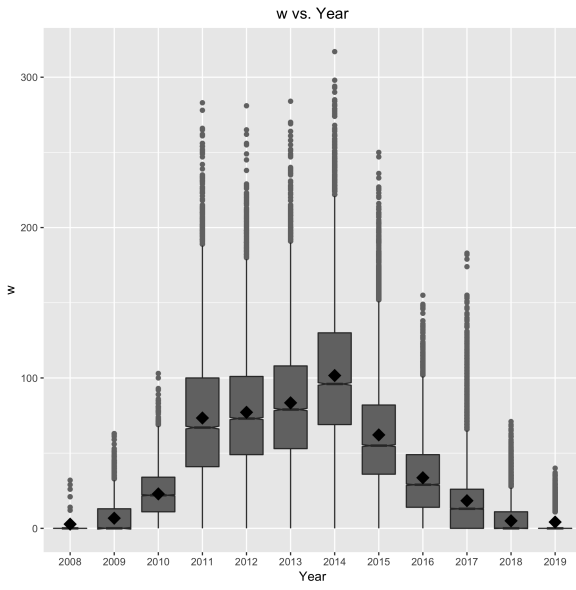


Figure 11: Box plots of raw Wolf number (w) by year.

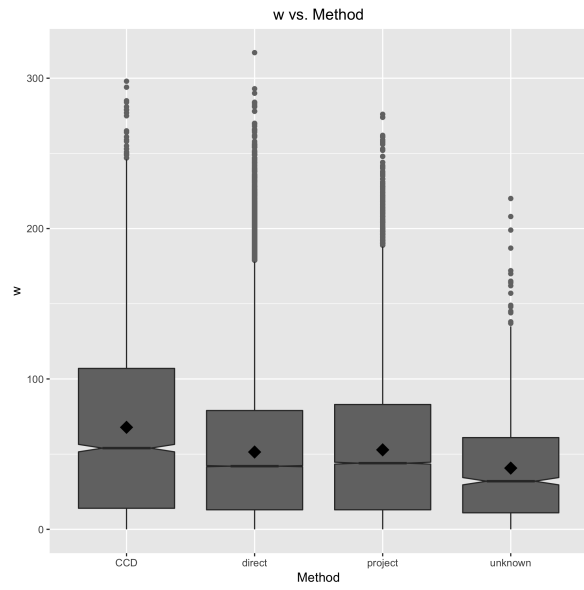


Figure 12: Box plots of raw Wolf number (w) by observing method.