

Monthly Report (00)
201708 Data Set

Wednesday 13th September, 2017

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.

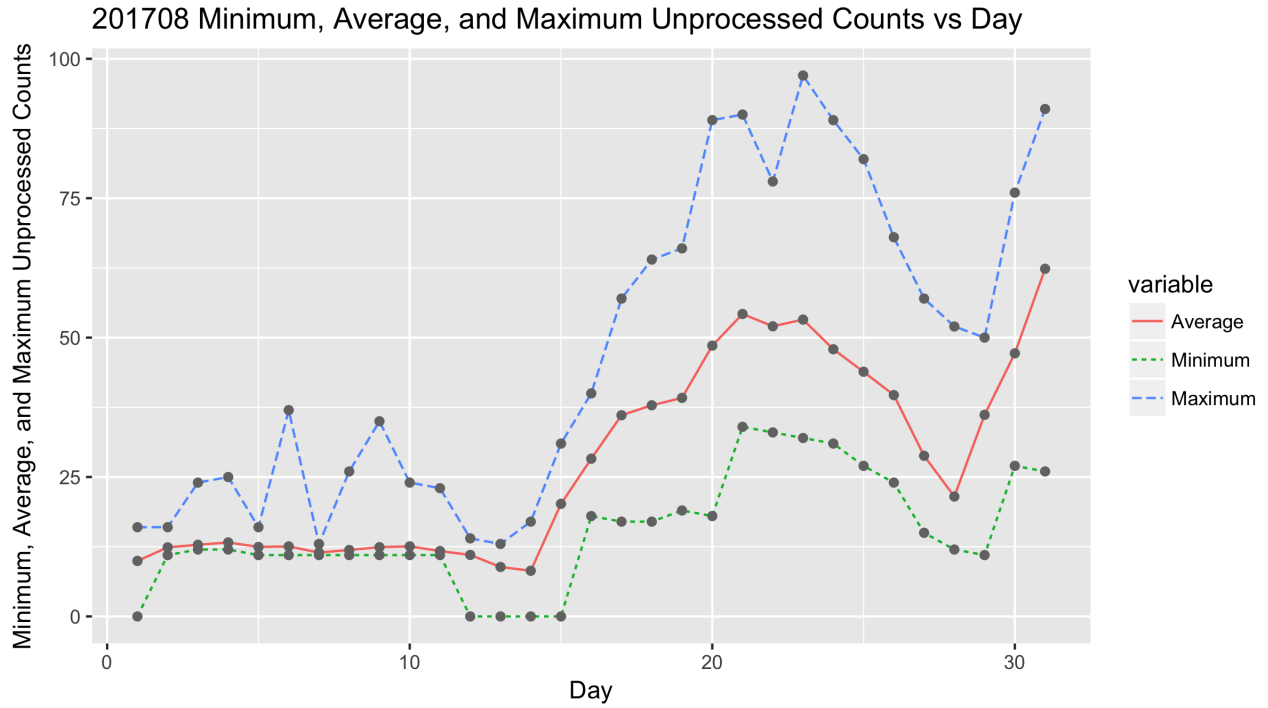


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201708 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	42.0000	0.0000	9.9524	16.0000
2.0000	38.0000	11.0000	12.3947	16.0000
3.0000	34.0000	12.0000	12.8529	24.0000
4.0000	37.0000	12.0000	13.2703	25.0000
5.0000	43.0000	11.0000	12.4651	16.0000
6.0000	36.0000	11.0000	12.5556	37.0000
7.0000	35.0000	11.0000	11.4571	13.0000
8.0000	36.0000	11.0000	11.9167	26.0000
9.0000	38.0000	11.0000	12.4211	35.0000
10.0000	37.0000	11.0000	12.5676	24.0000
11.0000	40.0000	11.0000	11.7250	23.0000
12.0000	35.0000	0.0000	11.0286	14.0000
13.0000	38.0000	0.0000	8.8684	13.0000
14.0000	37.0000	0.0000	8.1892	17.0000
15.0000	36.0000	0.0000	20.1944	31.0000
16.0000	31.0000	18.0000	28.2903	40.0000
17.0000	39.0000	17.0000	36.0769	57.0000
18.0000	37.0000	17.0000	37.8649	64.0000
19.0000	31.0000	19.0000	39.1935	66.0000
20.0000	36.0000	18.0000	48.5556	89.0000
21.0000	38.0000	34.0000	54.2368	90.0000
22.0000	32.0000	33.0000	52.0312	78.0000
23.0000	39.0000	32.0000	53.2308	97.0000
24.0000	35.0000	31.0000	47.9143	89.0000
25.0000	41.0000	27.0000	43.8780	82.0000
26.0000	36.0000	24.0000	39.6944	68.0000
27.0000	37.0000	15.0000	28.8108	57.0000
28.0000	30.0000	12.0000	21.5000	52.0000
29.0000	34.0000	11.0000	36.1471	50.0000
30.0000	33.0000	27.0000	47.1818	76.0000
31.0000	32.0000	26.0000	62.3438	91.0000

3 Error Tables

Data are for the month of August 2017. 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	silso
2010.05	23.6625	23.1219	24.2030	8.4000	8.7000
2010.06	18.3730	17.8830	18.8630	11.0000	13.6000
2010.07	20.2822	19.8341	20.7303	15.2000	16.1000
2010.08	20.7094	20.2060	21.2129	18.3000	19.6000
2010.09	23.9098	23.3969	24.4226	22.8000	25.2000
2010.10	22.7357	22.2461	23.2253	21.0000	23.5000
2010.11	23.4584	22.9304	23.9864	20.9000	21.6000
2010.12	22.4344	21.7888	23.0801	13.9000	14.5000
2011.01	76.0930	74.3819	77.8041	17.7000	18.7000
2011.02	66.4217	64.9281	67.9152	29.1000	29.6000
2011.03	71.2476	69.7769	72.7183	48.0000	55.8000
2011.04	78.5117	76.8349	80.1886	47.3000	54.4000
2011.05	79.8350	78.2461	81.4238	37.3000	41.5000
2011.06	65.2928	63.9211	66.6646	35.2000	37.0000
2011.07	70.2391	68.6941	71.7841	41.5000	43.8000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	silso
2011.08	75.1629	73.6913	76.6345	42.4000	50.5000
2011.09	83.8918	82.7900	84.9936	73.8000	78.0000
2011.10	79.1073	77.7526	80.4621	78.9000	88.0000
2011.11	80.3376	78.6147	82.0605	84.6000	96.7000
2011.12	74.7100	73.0621	76.3579	65.8000	73.0000
2012.01	78.2268	76.6743	79.7793	55.8000	58.2000
2012.02	66.0648	64.6616	67.4681	29.2000	33.1000
2012.03	73.6840	72.3691	74.9988	53.1000	64.1000
2012.04	78.2676	75.9584	80.5768	51.4000	55.2000
2012.05	83.7218	82.2627	85.1809	61.8000	69.0000
2012.06	68.2970	67.0920	69.5020	59.7000	64.5000
2012.07	74.4233	73.1725	75.6741	64.2000	51.3000
2012.08	75.8653	74.5993	77.1314	57.7000	63.1000
2012.09	84.7958	83.3440	86.2476	57.7000	61.5000
2012.10	81.5512	80.0097	83.0928	48.3000	53.3000
2012.11	83.8758	82.1933	85.5583	56.7000	61.4000
2012.12	75.7411	74.1483	77.3340	37.4000	40.8000
2013.01	88.0874	86.4431	89.7317	63.8000	62.9000
2013.02	76.1709	74.6869	77.6549	37.8000	38.0000
2013.03	80.9225	79.4017	82.4432	50.6000	57.9000
2013.04	91.0939	89.5604	92.6273	70.6000	72.4000
2013.05	91.3501	89.7775	92.9226	77.4000	78.7000
2013.06	75.1485	73.8178	76.4793	51.0000	52.5000
2013.07	79.9170	78.6733	81.1607	57.0000	57.0000
2013.08	83.6308	82.3244	84.9372	60.0000	66.0000
2013.09	92.5481	90.9475	94.1486	34.6000	36.9000
2013.10	87.4451	85.8937	88.9964	74.5000	85.6000
2013.11	90.0617	88.1903	91.9332	73.9000	77.6000
2013.12	83.3823	81.7074	85.0573	77.8000	90.3000
2014.01	104.6323	102.4343	106.8302	77.4000	82.0000
2014.02	90.4136	88.6967	92.1305	93.9000	102.8000
2014.03	99.5112	97.8338	101.1887	80.9000	92.2000
2014.04	110.7272	108.8476	112.6069	76.9000	84.7000
2014.05	110.5785	108.8084	112.3486	72.3000	75.2000
2014.06	90.9842	89.5090	92.4594	67.2000	71.0000
2014.07	98.0583	96.4597	99.6568	72.5000	72.5000
2014.08	102.0987	100.5648	103.6325	71.2000	74.7000
2014.09	114.1839	112.3400	116.0277	83.2000	87.6000
2014.10	107.6729	105.8726	109.4731	59.5000	60.6000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	silso
2014.11	111.4073	109.3118	113.5029	65.8000	71.1000
2014.12	100.8466	98.6569	103.0363	75.8000	78.0000
2015.01	63.8532	62.6248	65.0816	65.9000	67.0000
2015.02	55.1644	53.8774	56.4514	42.4000	44.8000
2015.03	59.5004	58.4088	60.5920	38.0000	38.4000
2015.04	67.0109	65.8226	68.1992	49.0000	54.4000
2015.05	66.5309	65.4622	67.5996	56.3000	58.8000
2015.06	55.2191	54.2782	56.1600	50.2000	68.3000
2015.07	58.3413	57.3443	59.3382	47.9000	65.8000
2015.08	62.2530	61.2323	63.2738	39.5000	57.2000
2015.09	69.3162	68.1870	70.4454	49.2000	72.1000
2015.10	65.3357	64.2218	66.4496	39.3000	48.3000
2015.11	68.1027	67.1815	69.0240	39.6000	55.9000
2015.12	61.1766	59.9460	62.4071	36.4000	44.8000
2016.01	35.6212	34.9885	36.2540	33.7000	43.3000
2016.02	30.1414	29.5438	30.7390	38.3000	46.8000
2016.03	32.3955	31.7999	32.9912	30.5000	38.9000
2016.04	35.8686	35.2321	36.5052	26.6000	30.9000
2016.05	36.6105	35.9848	37.2361	33.7000	48.4000
2016.06	29.9946	29.5192	30.4699	13.1000	19.5000
2016.07	32.3581	31.8585	32.8576	21.2000	27.5000
2016.08	34.1794	33.6167	34.7420	33.0000	47.9000
2016.09	37.8063	37.1792	38.4335	27.7000	37.1000
2016.10	35.8797	35.2594	36.4999	22.7000	31.7000
2016.11	36.9864	36.3501	37.6226	14.0000	22.2000
2016.12	33.5303	32.8589	34.2018	11.1000	20.0000
2017.01	21.3995	20.9925	21.8066	18.4000	26.2000
2017.02	17.9408	17.5891	18.2924	14.4000	20.6000
2017.03	19.7198	19.3712	20.0685	11.3000	15.5000
2017.04	22.2759	21.9152	22.6366	21.6000	33.2000
2017.05	22.1545	21.7948	22.5141	12.5000	18.1000
2017.06	18.0234	17.7617	18.2851	15.5000	19.3000
2017.07	19.6140	19.3446	19.8834	11.5000	16.3000
2017.08	20.5705	20.2671	20.8739	22.8000	35.7000

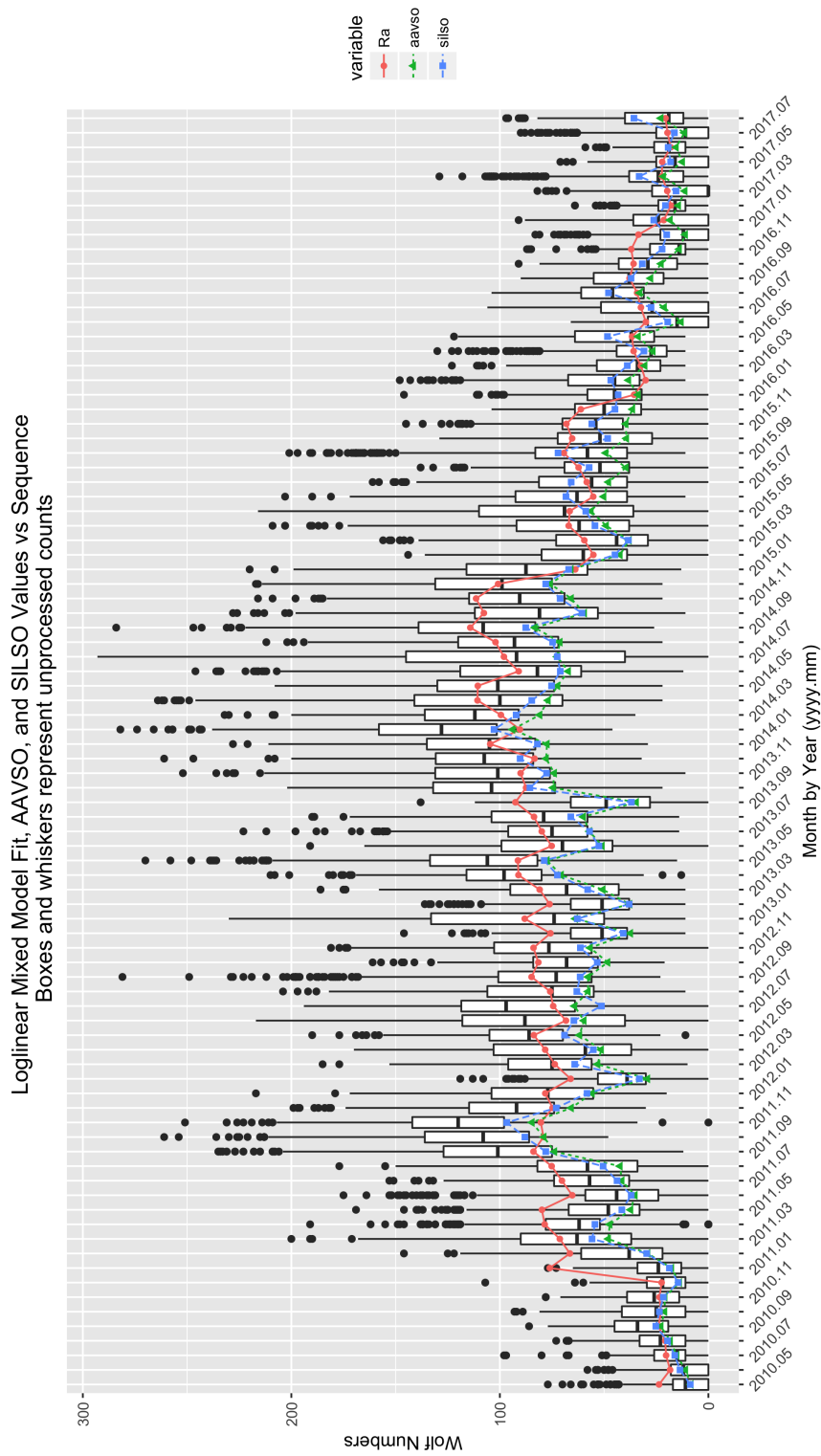


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201708 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2463	0.0327	99.3162	0.0000
seeF	-0.1897	0.0071	-26.5969	0.0000
seeG	-0.1015	0.0062	-16.3964	0.0000
seeP	-0.2936	0.0105	-28.0587	0.0000
silso1	0.0187	0.0389	0.4806	0.6308
year2011	1.2050	0.0154	78.1136	0.0000
year2012	1.2234	0.0154	79.5702	0.0000
year2013	1.3182	0.0153	85.9856	0.0000
year2014	1.5068	0.0152	98.9472	0.0000
year2015	1.0087	0.0156	64.5609	0.0000
year2016	0.4039	0.0166	24.3078	0.0000
year2017	-0.1041	0.0191	-5.4631	0.0000
mon2	-0.1554	0.0119	-13.1163	0.0000
mon3	-0.0790	0.0109	-7.2276	0.0000
mon4	0.0308	0.0109	2.8240	0.0047
mon5	0.0335	0.0104	3.2227	0.0013
mon6	-0.1712	0.0109	-15.6840	0.0000
mon7	-0.1015	0.0105	-9.6376	0.0000
mon8	-0.0500	0.0103	-4.8506	0.0000
mon9	0.0616	0.0101	6.1090	0.0000
mon10	0.0090	0.0107	0.8475	0.3967
mon11	0.0473	0.0109	4.3307	0.0000
mon12	-0.0421	0.0116	-3.6205	0.0003

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 (silso) 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: 201708 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2615	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2401	1st Qu.:2456080	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2393	Median :2456706	Median :2014	Median : 7.000	Median :16.00
BROB : 2126	Mean :2456424	Mean :2014	Mean : 6.604	Mean :15.72
KNJS : 2012	3rd Qu.:2457347	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 1990	Max. :2457997	Max. :2017	Max. :12.000	Max. :31.00
(Other):48642				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:11535	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24894	Min. :0.0000
F:19010	1st Qu.: 2.000	1st Qu.: 7.00	1st Qu.: 29.00	3000F : 9764	1st Qu.:0.0000
G:26589	Median : 4.000	Median : 19.00	Median : 58.00	2500E : 7766	Median :0.0000
P: 5045	Mean : 3.993	Mean : 24.86	Mean : 64.79	3500G : 4618	Mean :0.3277
	3rd Qu.: 6.000	3rd Qu.: 36.00	3rd Qu.: 93.00	1000B : 4228	3rd Qu.:1.0000
	Max. :18.000	Max. :204.00	Max. :293.00	: 3503	Max. :1.0000
				(Other): 7406	

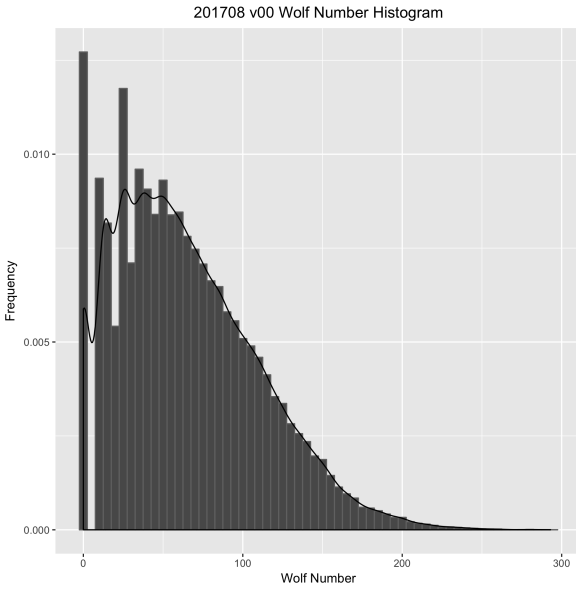


Figure 3: Box plots of raw Wolf number (w) by observer rank.

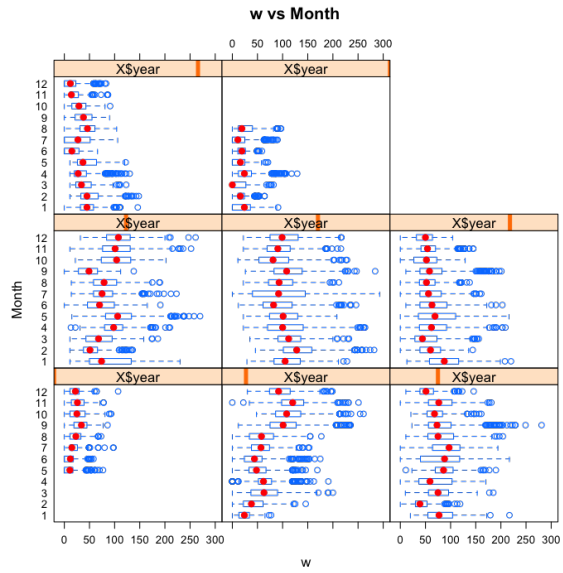


Figure 4: Box plots of raw Wolf number (w) by month and year.

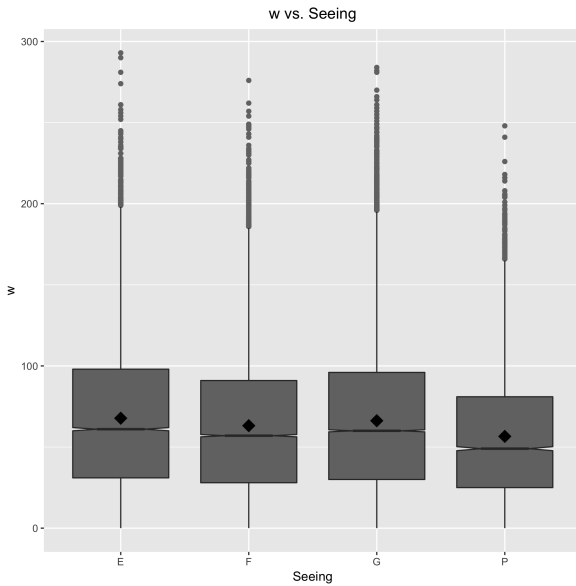


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

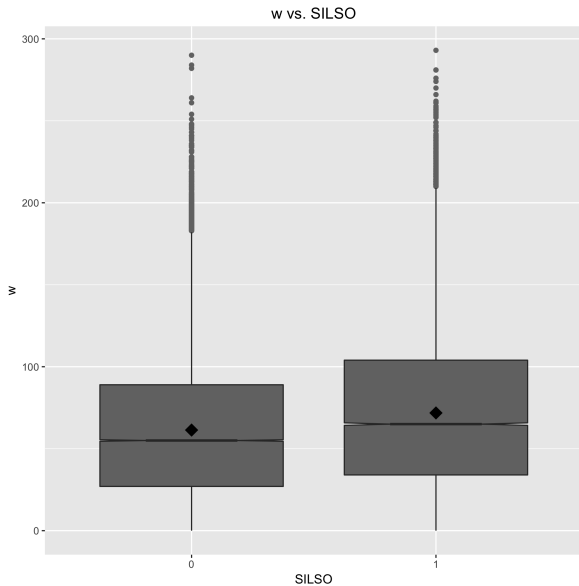


Figure 6: Box plots of raw Wolf number (w) by organization.

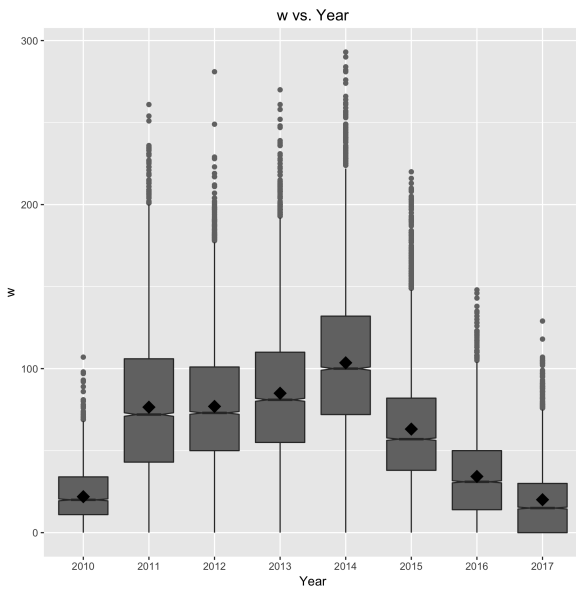


Figure 7: Box plots of raw Wolf number (w) by year.