

Monthly Report (00)
201709 Data Set

Saturday 14th October, 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.

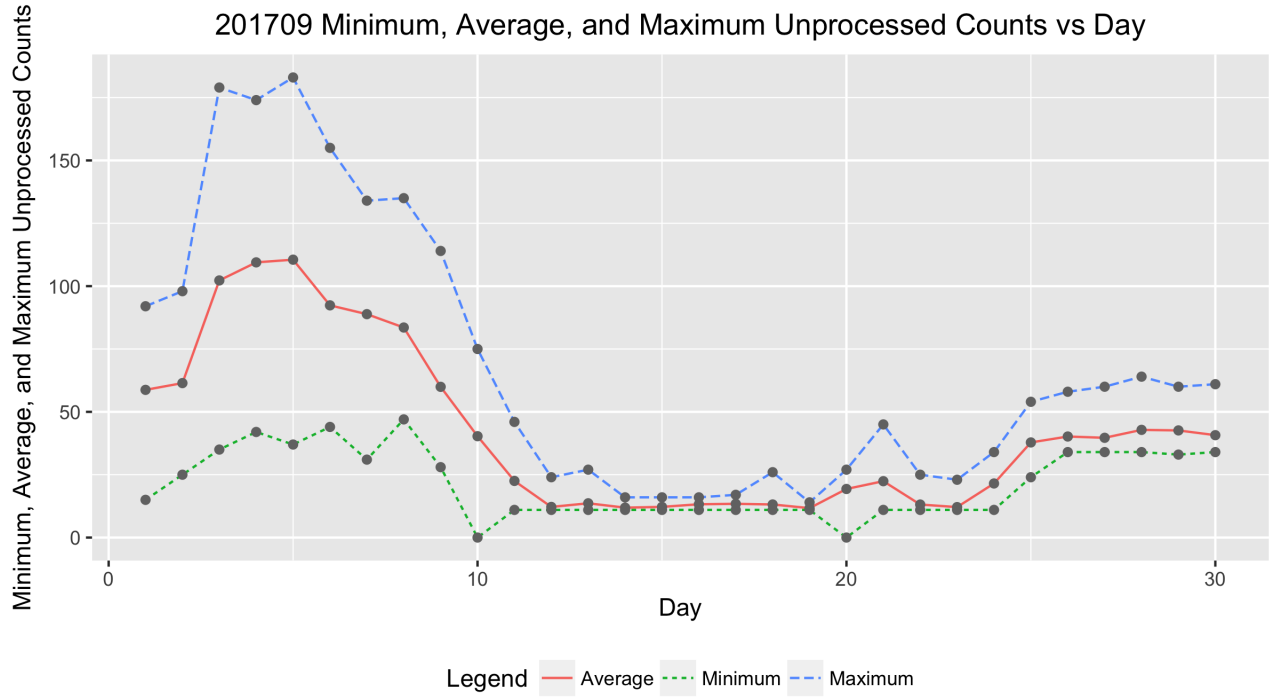


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201709 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	39.0000	15.0000	58.7436	92.0000
2.0000	39.0000	25.0000	61.4359	98.0000
3.0000	41.0000	35.0000	102.2683	179.0000
4.0000	37.0000	42.0000	109.4595	174.0000
5.0000	36.0000	37.0000	110.5278	183.0000
6.0000	39.0000	44.0000	92.3590	155.0000
7.0000	36.0000	31.0000	88.8611	134.0000
8.0000	35.0000	47.0000	83.5429	135.0000
9.0000	41.0000	28.0000	59.9512	114.0000
10.0000	38.0000	0.0000	40.3158	75.0000
11.0000	36.0000	11.0000	22.5278	46.0000
12.0000	33.0000	11.0000	12.1818	24.0000
13.0000	40.0000	11.0000	13.6250	27.0000
14.0000	37.0000	11.0000	11.8919	16.0000
15.0000	40.0000	11.0000	12.2000	16.0000
16.0000	37.0000	11.0000	13.2432	16.0000
17.0000	38.0000	11.0000	13.4211	17.0000
18.0000	38.0000	11.0000	13.1579	26.0000
19.0000	40.0000	11.0000	11.7000	14.0000
20.0000	37.0000	0.0000	19.3243	27.0000
21.0000	40.0000	11.0000	22.4250	45.0000
22.0000	37.0000	11.0000	13.1351	25.0000
23.0000	35.0000	11.0000	12.1143	23.0000
24.0000	41.0000	11.0000	21.5122	34.0000
25.0000	36.0000	24.0000	37.8333	54.0000
26.0000	36.0000	34.0000	40.1944	58.0000
27.0000	34.0000	34.0000	39.6765	60.0000
28.0000	30.0000	34.0000	42.8333	64.0000
29.0000	31.0000	33.0000	42.6129	60.0000
30.0000	29.0000	34.0000	40.7241	61.0000

3 Error Tables

Data are for the month of September 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	23.5047	22.9679	24.0416	8.4000	8.7000
2010.06	18.1755	17.6901	18.6609	11.0000	13.6000
2010.07	20.0688	19.6249	20.5127	15.2000	16.1000
2010.08	20.5164	20.0173	21.0156	18.3000	19.6000
2010.09	24.6961	24.1663	25.2260	22.8000	25.2000
2010.10	22.6945	22.2049	23.1841	21.0000	23.5000
2010.11	23.4215	22.8937	23.9493	20.9000	21.6000
2010.12	22.4092	21.7647	23.0537	13.9000	14.5000
2011.01	75.3468	73.6536	77.0399	17.7000	18.7000
2011.02	65.8041	64.3281	67.2800	29.1000	29.6000
2011.03	70.5782	69.1218	72.0345	48.0000	55.8000
2011.04	77.7269	76.0668	79.3871	47.3000	54.4000
2011.05	79.1017	77.5311	80.6723	37.3000	41.5000
2011.06	64.4321	63.0805	65.7837	35.2000	37.0000
2011.07	69.3417	67.8175	70.8659	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	74.2766	72.8261	75.7271	42.4000	50.5000
2011.09	86.3382	85.2050	87.4715	73.8000	78.0000
2011.10	78.7513	77.4096	80.0931	78.9000	88.0000
2011.11	79.9438	78.2330	81.6547	84.6000	96.7000
2011.12	74.3596	72.7210	75.9982	65.8000	73.0000
2012.01	77.8560	76.3140	79.3980	55.8000	58.2000
2012.02	65.7646	64.3710	67.1582	29.2000	33.1000
2012.03	73.3397	72.0339	74.6456	53.1000	64.1000
2012.04	77.8764	75.5745	80.1783	51.4000	55.2000
2012.05	83.3370	81.8889	84.7851	61.8000	69.0000
2012.06	67.7117	66.5223	68.9012	59.7000	64.5000
2012.07	73.8280	72.5908	75.0653	64.2000	51.3000
2012.08	75.3487	74.0950	76.6023	57.7000	63.1000
2012.09	87.8065	86.3056	89.3075	57.7000	61.5000
2012.10	81.6056	80.0647	83.1466	48.3000	53.3000
2012.11	83.9331	82.2494	85.6167	56.7000	61.4000
2012.12	75.8102	74.2187	77.4017	37.4000	40.8000
2013.01	87.7012	86.0668	89.3356	63.8000	62.9000
2013.02	75.8654	74.3903	77.3404	37.8000	38.0000
2013.03	80.6101	79.0992	82.1210	50.6000	57.9000
2013.04	90.6485	89.1247	92.1724	70.6000	72.4000
2013.05	90.9676	89.4052	92.5299	77.4000	78.7000
2013.06	74.5248	73.2093	75.8402	51.0000	52.5000
2013.07	79.3218	78.0905	80.5531	57.0000	57.0000
2013.08	83.0993	81.8030	84.3955	60.0000	66.0000
2013.09	95.8528	94.1962	97.5093	34.6000	36.9000
2013.10	87.5431	85.9911	89.0950	74.5000	85.6000
2013.11	90.1831	88.3117	92.0545	73.9000	77.6000
2013.12	83.4679	81.7908	85.1451	77.8000	90.3000
2014.01	104.1989	102.0137	106.3840	77.4000	82.0000
2014.02	90.0602	88.3516	91.7689	93.9000	102.8000
2014.03	99.1100	97.4424	100.7777	80.9000	92.2000
2014.04	110.1828	108.3156	112.0500	76.9000	84.7000
2014.05	110.1330	108.3732	111.8928	72.3000	75.2000
2014.06	90.2459	88.7848	91.7070	67.2000	71.0000
2014.07	97.3101	95.7246	98.8955	72.5000	72.5000
2014.08	101.4297	99.9087	102.9506	71.2000	74.7000
2014.09	118.2322	116.3246	120.1399	83.2000	87.6000
2014.10	107.8083	106.0095	109.6071	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	111.5241	109.4308	113.6174	65.8000	71.1000
2014.12	100.9624	98.7734	103.1514	75.8000	78.0000
2015.01	63.5859	62.3680	64.8038	65.9000	67.0000
2015.02	54.9199	53.6421	56.1976	42.4000	44.8000
2015.03	59.2562	58.1722	60.3403	38.0000	38.4000
2015.04	66.6526	65.4726	67.8327	49.0000	54.4000
2015.05	66.2287	65.1659	67.2914	56.3000	58.8000
2015.06	54.7556	53.8246	55.6866	50.2000	68.3000
2015.07	57.8884	56.9010	58.8758	47.9000	65.8000
2015.08	61.8309	60.8199	62.8420	39.5000	57.2000
2015.09	71.7900	70.6224	72.9576	49.2000	72.1000
2015.10	65.3895	64.2767	66.5022	39.3000	48.3000
2015.11	68.1563	67.2361	69.0766	39.6000	55.9000
2015.12	61.2509	60.0227	62.4791	36.4000	44.8000
2016.01	35.4600	34.8311	36.0889	33.7000	43.3000
2016.02	30.0235	29.4299	30.6171	38.3000	46.8000
2016.03	32.2595	31.6672	32.8517	30.5000	38.9000
2016.04	35.6951	35.0632	36.3269	26.6000	30.9000
2016.05	36.4593	35.8380	37.0806	33.7000	48.4000
2016.06	29.7545	29.2843	30.2246	13.1000	19.5000
2016.07	32.1053	31.6106	32.5999	21.2000	27.5000
2016.08	33.9547	33.3972	34.5121	33.0000	47.9000
2016.09	39.1451	38.4967	39.7934	27.7000	37.1000
2016.10	35.9246	35.3044	36.5447	22.7000	31.7000
2016.11	37.0322	36.3960	37.6685	14.0000	22.2000
2016.12	33.5821	32.9114	34.2527	11.1000	20.0000
2017.01	23.8496	23.3967	24.3025	18.4000	26.2000
2017.02	20.0013	19.6097	20.3929	14.4000	20.6000
2017.03	21.9836	21.5959	22.3712	11.3000	15.5000
2017.04	24.7785	24.3759	25.1812	21.6000	33.2000
2017.05	24.6707	24.2701	25.0713	12.5000	18.1000
2017.06	20.3271	20.0138	20.6404	15.5000	19.3000
2017.07	22.0954	21.7704	22.4204	11.5000	16.3000
2017.08	23.1642	22.7978	23.5306	22.8000	35.7000
2017.09	27.3167	26.8544	27.7790	34.6000	42.9000

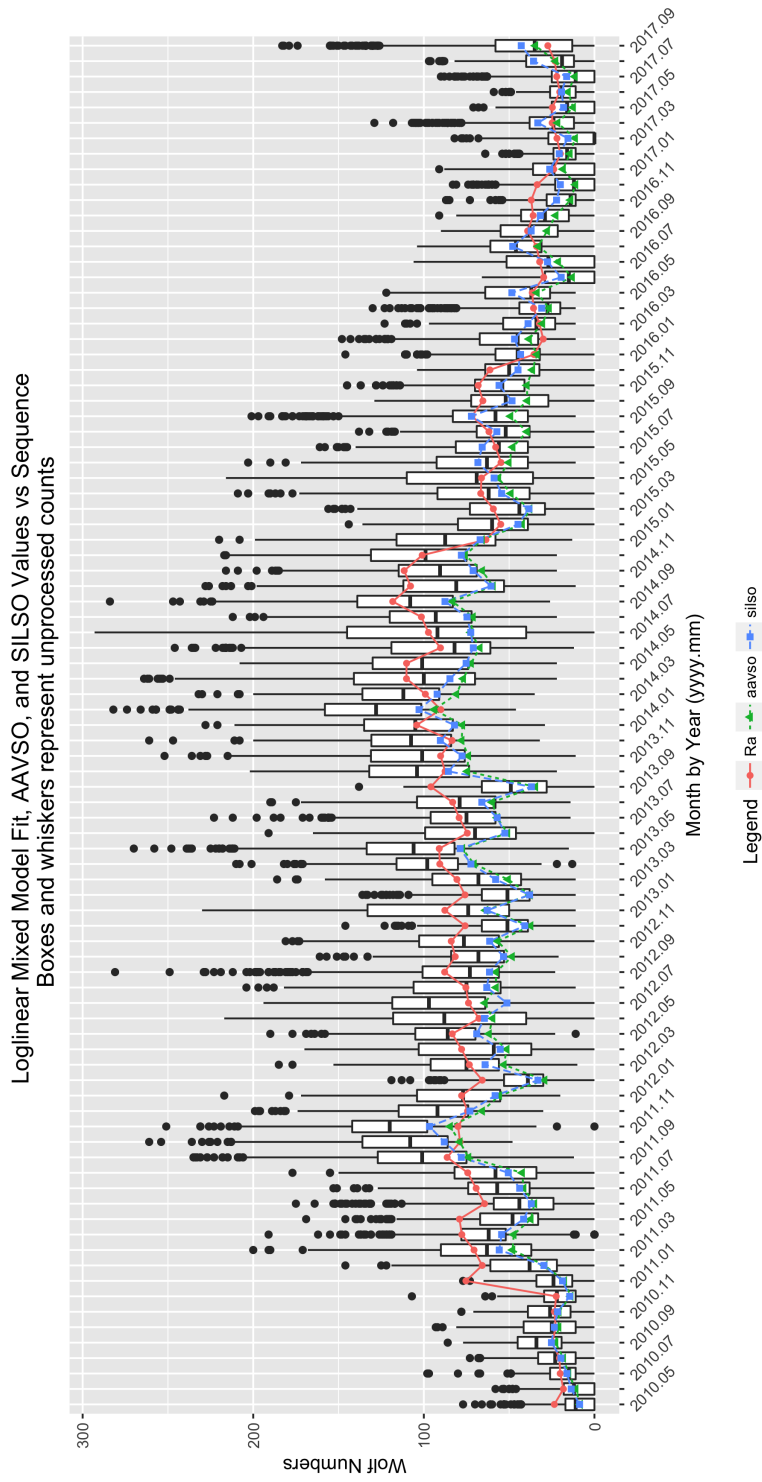


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201709 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2691	0.0343	95.3119	0.0000
seeF	-0.1920	0.0072	-26.8327	0.0000
seeG	-0.1032	0.0062	-16.6033	0.0000
seeP	-0.2960	0.0105	-28.1765	0.0000
silso1	-0.0257	0.0353	-0.7289	0.4661
year2011	1.2023	0.0156	77.2610	0.0000
year2012	1.2260	0.0155	79.0532	0.0000
year2013	1.3212	0.0155	85.4404	0.0000
year2014	1.5099	0.0154	98.2959	0.0000
year2015	1.0116	0.0158	64.1850	0.0000
year2016	0.4070	0.0168	24.2814	0.0000
year2017	0.0118	0.0183	0.6426	0.5205
mon2	-0.1551	0.0120	-12.9762	0.0000
mon3	-0.0787	0.0110	-7.1423	0.0000
mon4	0.0302	0.0110	2.7417	0.0061
mon5	0.0335	0.0105	3.1996	0.0014
mon6	-0.1755	0.0110	-15.9393	0.0000
mon7	-0.1052	0.0106	-9.9106	0.0000
mon8	-0.0525	0.0104	-5.0518	0.0000
mon9	0.1009	0.0100	10.0780	0.0000
mon10	0.0144	0.0108	1.3427	0.1794
mon11	0.0528	0.0110	4.7916	0.0000
mon12	-0.0364	0.0117	-3.1106	0.0019

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: 201709 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2645	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2429	1st Qu.:2456090	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2418	Median :2456730	Median :2014	Median : 7.000	Median :16.00
BROB : 2156	Mean :2456451	Mean :2014	Mean : 6.645	Mean :15.71
KNJS : 2042	3rd Qu.:2457383	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 2015	Max. :2458027	Max. :2017	Max. :12.000	Max. :31.00
(Other):49551				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:11778	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24894	Min. :0.0000
F:19297	1st Qu.: 2.000	1st Qu.: 7.00	1st Qu.: 29.00	3000F : 9764	1st Qu.:0.0000
G:27065	Median : 4.000	Median : 18.00	Median : 57.00	2500E : 7766	Median :0.0000
P: 5116	Mean : 3.968	Mean : 24.73	Mean : 64.41	3500G : 4618	Mean :0.3274
	3rd Qu.: 6.000	3rd Qu.: 36.00	3rd Qu.: 93.00	: 4580	3rd Qu.:1.0000
	Max. :18.000	Max. :204.00	Max. :293.00	1000B : 4228	Max. :1.0000
				(Other): 7406	

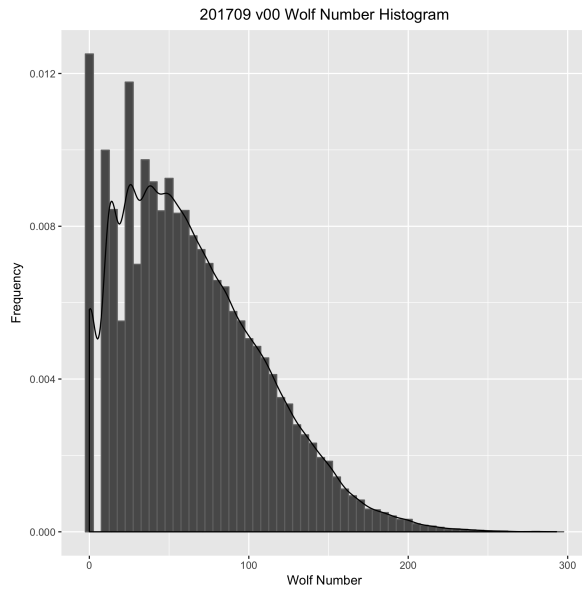


Figure 3: Box plots of raw Wolf number (w) by observer rank.

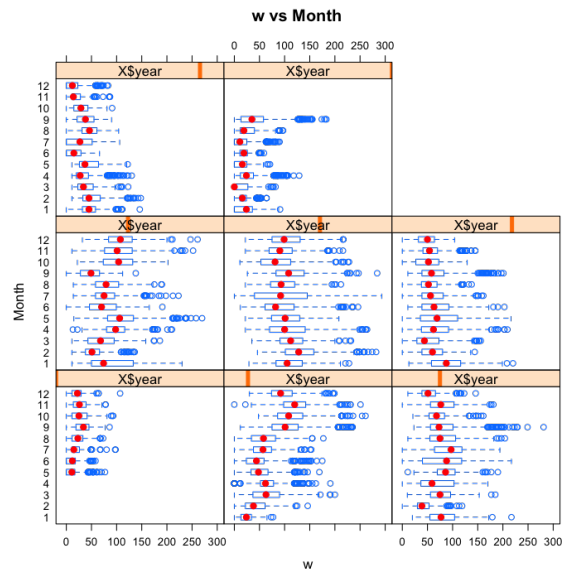


Figure 4: Box plots of raw Wolf number (w) by month and year.

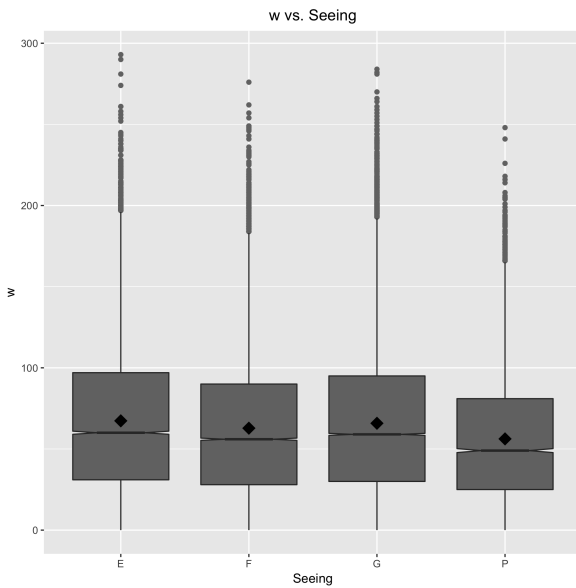


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

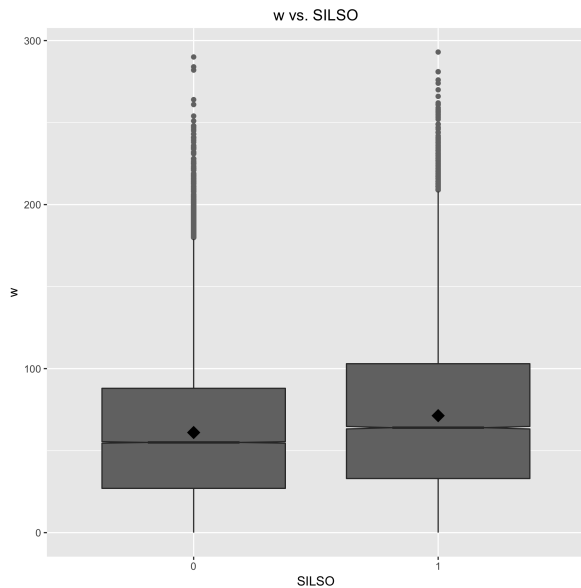


Figure 6: Box plots of raw Wolf number (w) by organization.

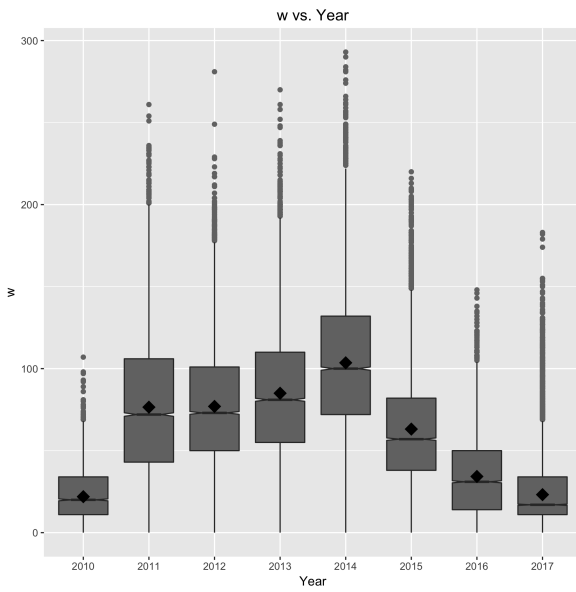


Figure 7: Box plots of raw Wolf number (w) by year.