

Monthly Report (00)

201711 Data Set

Tuesday 12th December, 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.

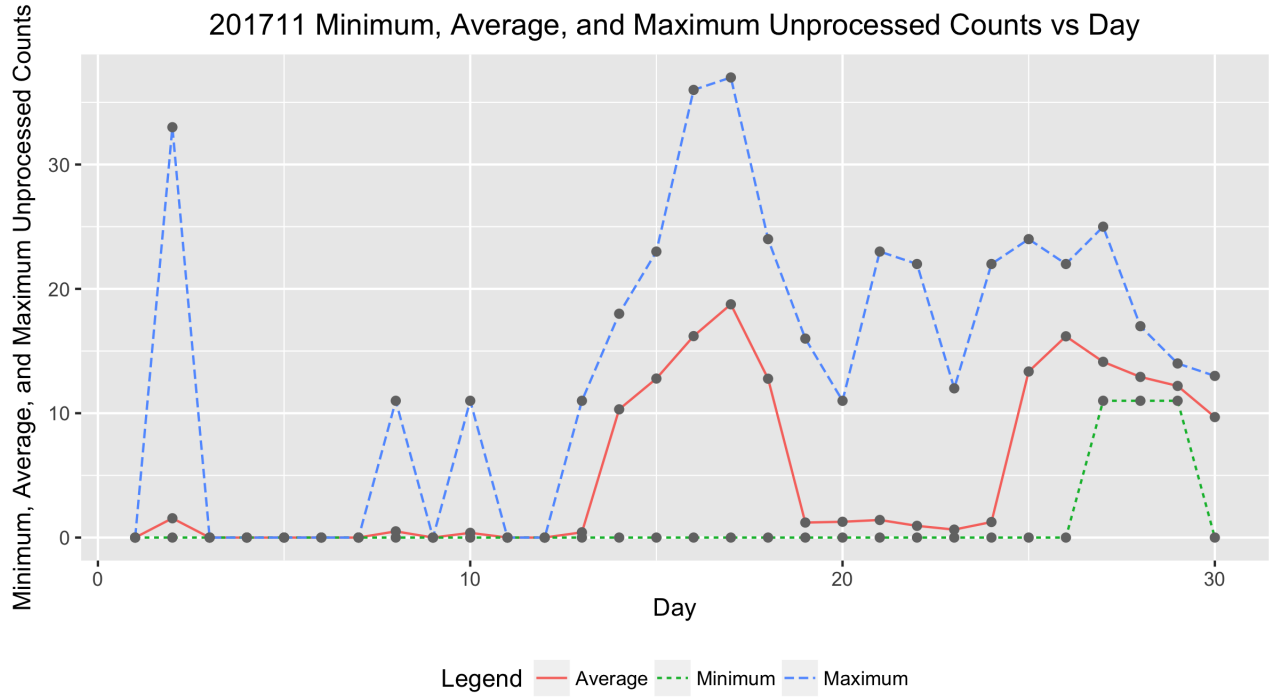


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201711 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	30.0000	0.0000	0.0000	0.0000
2.0000	29.0000	0.0000	1.5517	33.0000
3.0000	30.0000	0.0000	0.0000	0.0000
4.0000	25.0000	0.0000	0.0000	0.0000
5.0000	30.0000	0.0000	0.0000	0.0000
6.0000	25.0000	0.0000	0.0000	0.0000
7.0000	26.0000	0.0000	0.0000	0.0000
8.0000	22.0000	0.0000	0.5000	11.0000
9.0000	32.0000	0.0000	0.0000	0.0000
10.0000	29.0000	0.0000	0.3793	11.0000
11.0000	34.0000	0.0000	0.0000	0.0000
12.0000	32.0000	0.0000	0.0000	0.0000
13.0000	26.0000	0.0000	0.4231	11.0000
14.0000	29.0000	0.0000	10.3103	18.0000
15.0000	33.0000	0.0000	12.7879	23.0000
16.0000	30.0000	0.0000	16.2000	36.0000
17.0000	33.0000	0.0000	18.7576	37.0000
18.0000	31.0000	0.0000	12.7742	24.0000
19.0000	34.0000	0.0000	1.2059	16.0000
20.0000	26.0000	0.0000	1.2692	11.0000
21.0000	29.0000	0.0000	1.4138	23.0000
22.0000	35.0000	0.0000	0.9429	22.0000
23.0000	36.0000	0.0000	0.6389	12.0000
24.0000	36.0000	0.0000	1.2500	22.0000
25.0000	34.0000	0.0000	13.3529	24.0000
26.0000	38.0000	0.0000	16.1842	22.0000
27.0000	30.0000	11.0000	14.1333	25.0000
28.0000	37.0000	11.0000	12.9189	17.0000
29.0000	36.0000	11.0000	12.1944	14.0000
30.0000	29.0000	0.0000	9.6897	13.0000

3 Error Tables

Data are for the month of November 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	23.7160	23.1730	24.2589	8.4000	8.7000
2010.06	18.4262	17.9335	18.9189	11.0000	13.6000
2010.07	20.3367	19.8863	20.7870	15.2000	16.1000
2010.08	20.7714	20.2649	21.2779	18.3000	19.6000
2010.09	24.9840	24.4470	25.5210	22.8000	25.2000
2010.10	22.2542	21.7744	22.7340	21.0000	23.5000
2010.11	22.4710	21.9639	22.9781	20.9000	21.6000
2010.12	22.4856	21.8384	23.1329	13.9000	14.5000
2011.01	75.7549	74.0499	77.4599	17.7000	18.7000
2011.02	66.1311	64.6458	67.6163	29.1000	29.6000
2011.03	70.9363	69.4710	72.4016	48.0000	55.8000
2011.04	78.1843	76.5149	79.8538	47.3000	54.4000
2011.05	79.4755	77.8996	81.0513	37.3000	41.5000
2011.06	65.0541	63.6916	66.4166	35.2000	37.0000
2011.07	69.9821	68.4430	71.5212	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	74.9149	73.4477	76.3820	42.4000	50.5000
2011.09	86.9722	85.8266	88.1178	73.8000	78.0000
2011.10	76.9235	75.6096	78.2375	78.9000	88.0000
2011.11	76.4335	74.7947	78.0724	84.6000	96.7000
2011.12	74.3689	72.7276	76.0101	65.8000	73.0000
2012.01	78.2852	76.7339	79.8364	55.8000	58.2000
2012.02	66.1071	64.7019	67.5123	29.2000	33.1000
2012.03	73.7156	72.3987	75.0325	53.1000	64.1000
2012.04	78.3846	76.0621	80.7070	51.4000	55.2000
2012.05	83.7652	82.3068	85.2236	61.8000	69.0000
2012.06	68.3980	67.1925	69.6035	59.7000	64.5000
2012.07	74.5135	73.2602	75.7667	64.2000	51.3000
2012.08	75.9451	74.6757	77.2144	57.7000	63.1000
2012.09	88.4784	86.9643	89.9925	57.7000	61.5000
2012.10	79.7178	78.2128	81.2227	48.3000	53.3000
2012.11	80.2451	78.6356	81.8546	56.7000	61.4000
2012.12	75.8005	74.2092	77.3919	37.4000	40.8000
2013.01	88.1312	86.4902	89.7722	63.8000	62.9000
2013.02	76.2169	74.7348	77.6989	37.8000	38.0000
2013.03	80.9792	79.4588	82.4997	50.6000	57.9000
2013.04	91.1480	89.6154	92.6805	70.6000	72.4000
2013.05	91.4061	89.8337	92.9785	77.4000	78.7000
2013.06	75.2220	73.8900	76.5541	51.0000	52.5000
2013.07	80.0069	78.7628	81.2510	57.0000	57.0000
2013.08	83.6982	82.3905	85.0058	60.0000	66.0000
2013.09	96.5242	94.8558	98.1927	34.6000	36.9000
2013.10	85.4089	83.8942	86.9237	74.5000	85.6000
2013.11	86.1527	84.3649	87.9406	73.9000	77.6000
2013.12	83.3605	81.6845	85.0365	77.8000	90.3000
2014.01	104.7171	102.5173	106.9169	77.4000	82.0000
2014.02	90.4814	88.7645	92.1983	93.9000	102.8000
2014.03	99.5646	97.8860	101.2433	80.9000	92.2000
2014.04	110.7890	108.9117	112.6664	76.9000	84.7000
2014.05	110.5869	108.8139	112.3598	72.3000	75.2000
2014.06	91.1369	89.6627	92.6110	67.2000	71.0000
2014.07	98.2208	96.6203	99.8213	72.5000	72.5000
2014.08	102.2088	100.6755	103.7421	71.2000	74.7000
2014.09	119.0846	117.1642	121.0050	83.2000	87.6000
2014.10	105.2489	103.4946	107.0032	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	106.5476	104.5467	108.5485	65.8000	71.1000
2014.12	100.9080	98.7210	103.0951	75.8000	78.0000
2015.01	64.0887	62.8578	65.3197	65.9000	67.0000
2015.02	55.3531	54.0622	56.6440	42.4000	44.8000
2015.03	59.7223	58.6281	60.8165	38.0000	38.4000
2015.04	67.2282	66.0372	68.4192	49.0000	54.4000
2015.05	66.7449	65.6738	67.8159	56.3000	58.8000
2015.06	55.4416	54.4973	56.3859	50.2000	68.3000
2015.07	58.5861	57.5852	59.5871	47.9000	65.8000
2015.08	62.5255	61.5039	63.5470	39.5000	57.2000
2015.09	72.5355	71.3553	73.7158	49.2000	72.1000
2015.10	64.0358	62.9463	65.1252	39.3000	48.3000
2015.11	65.3406	64.4590	66.2221	39.6000	55.9000
2015.12	61.4067	60.1767	62.6367	36.4000	44.8000
2016.01	35.6472	35.0153	36.2792	33.7000	43.3000
2016.02	30.1785	29.5817	30.7754	38.3000	46.8000
2016.03	32.4280	31.8321	33.0239	30.5000	38.9000
2016.04	35.9015	35.2651	36.5378	26.6000	30.9000
2016.05	36.6413	36.0170	37.2656	33.7000	48.4000
2016.06	30.0488	29.5737	30.5238	13.1000	19.5000
2016.07	32.4047	31.9046	32.9048	21.2000	27.5000
2016.08	34.2369	33.6748	34.7990	33.0000	47.9000
2016.09	39.4276	38.7737	40.0815	27.7000	37.1000
2016.10	35.0713	34.4652	35.6774	22.7000	31.7000
2016.11	35.3870	34.7789	35.9951	14.0000	22.2000
2016.12	33.5521	32.8806	34.2235	11.1000	20.0000
2017.01	21.2656	20.8617	21.6695	18.4000	26.2000
2017.02	17.8306	17.4814	18.1799	14.4000	20.6000
2017.03	19.5952	19.2491	19.9412	11.3000	15.5000
2017.04	22.1380	21.7805	22.4955	21.6000	33.2000
2017.05	22.0152	21.6583	22.3721	12.5000	18.1000
2017.06	17.7096	17.4319	17.9872	15.5000	19.3000
2017.07	19.2685	18.9829	19.5541	11.5000	16.3000
2017.08	20.2171	19.8966	20.5375	22.8000	35.7000
2017.09	23.7315	23.3269	24.1361	34.6000	42.9000
2017.10	20.3896	20.0308	20.7484	10.5000	11.0000
2017.11	20.2921	19.9082	20.6760	4.2000	5.6000

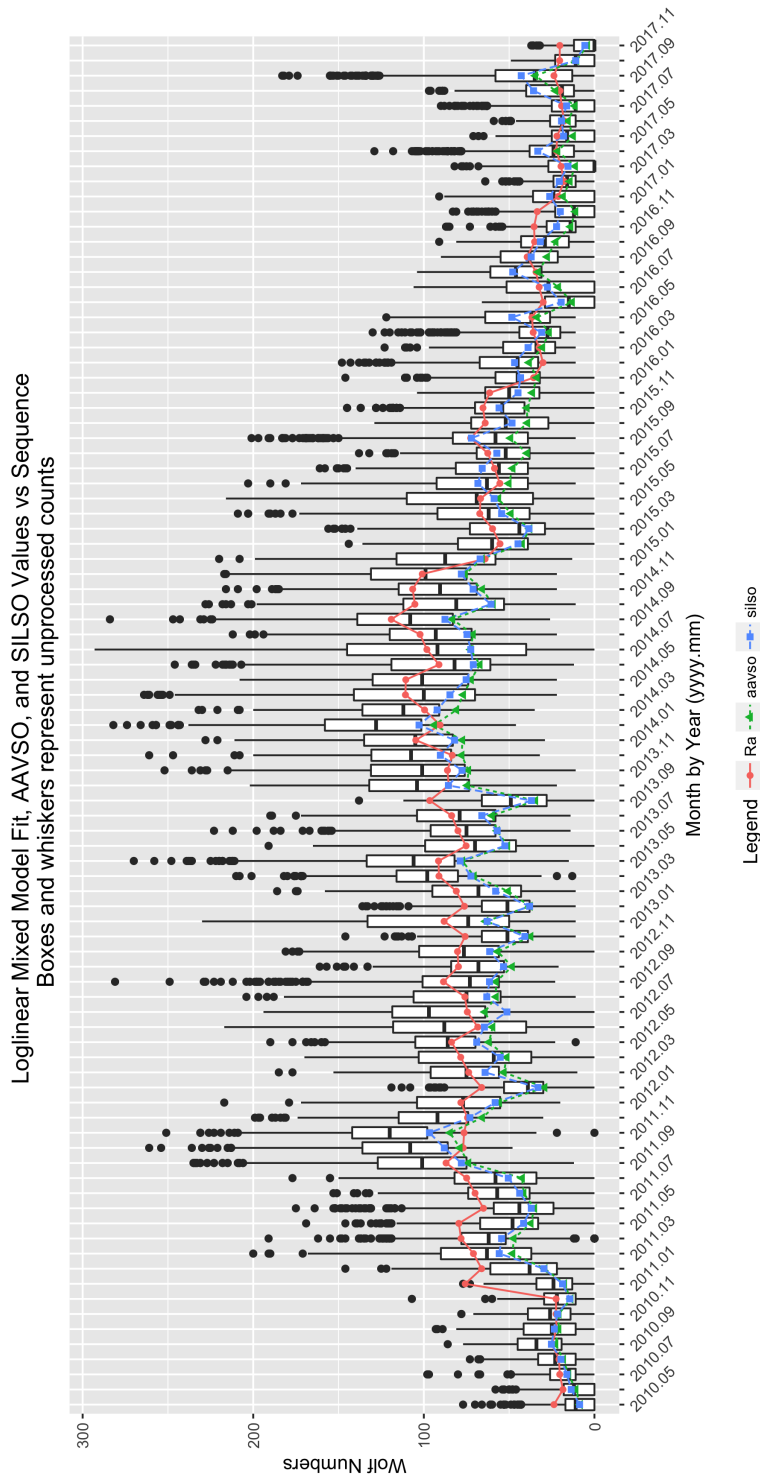


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201711 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2370	0.0352	91.9977	0.0000
seeF	-0.1908	0.0072	-26.5652	0.0000
seeG	-0.1024	0.0062	-16.4241	0.0000
seeP	-0.2931	0.0105	-27.8100	0.0000
silso1	0.0175	0.0335	0.5223	0.6015
year2011	1.1979	0.0157	76.5281	0.0000
year2012	1.2215	0.0156	78.2990	0.0000
year2013	1.3159	0.0156	84.5965	0.0000
year2014	1.5049	0.0155	97.3919	0.0000
year2015	1.0096	0.0159	63.6815	0.0000
year2016	0.4018	0.0169	23.8324	0.0000
year2017	-0.1133	0.0181	-6.2434	0.0000
mon2	-0.1552	0.0120	-12.9088	0.0000
mon3	-0.0785	0.0111	-7.0837	0.0000
mon4	0.0313	0.0111	2.8292	0.0047
mon5	0.0339	0.0105	3.2195	0.0013
mon6	-0.1702	0.0111	-15.3750	0.0000
mon7	-0.1003	0.0107	-9.3997	0.0000
mon8	-0.0490	0.0105	-4.6833	0.0000
mon9	0.1036	0.0101	10.2979	0.0000
mon10	-0.0144	0.0107	-1.3417	0.1797
mon11	0.0021	0.0110	0.1943	0.8459
mon12	-0.0424	0.0118	-3.5948	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: 201711 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2705	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2486	1st Qu.:2456106	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2431	Median :2456769	Median :2014	Median : 7.000	Median :16.00
BROB : 2183	Mean :2456499	Mean :2014	Mean : 6.759	Mean :15.71
KNJS : 2103	3rd Qu.:2457464	3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 2063	Max. :2458088	Max. :2017	Max. :12.000	Max. :31.00
(Other):51223				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:12199	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24894	Min. :0.0000
F:19886	1st Qu.: 2.000	1st Qu.: 6.00	1st Qu.: 26.00	3000F : 9764	1st Qu.:0.0000
G:27833	Median : 4.000	Median : 18.00	Median : 56.00	2500E : 7766	Median :0.0000
P: 5276	Mean : 3.872	Mean : 24.05	Mean : 62.77	: 6518	Mean :0.3277
	3rd Qu.: 6.000	3rd Qu.: 35.00	3rd Qu.: 92.00	3500G : 4618	3rd Qu.:1.0000
	Max. :18.000	Max. :204.00	Max. :293.00	1000B : 4228	Max. :1.0000
				(Other): 7406	

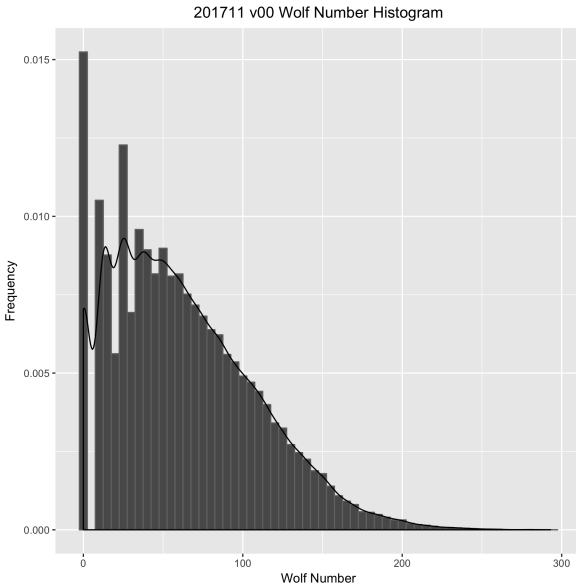


Figure 3: Box plots of raw Wolf number (w) by observer rank.

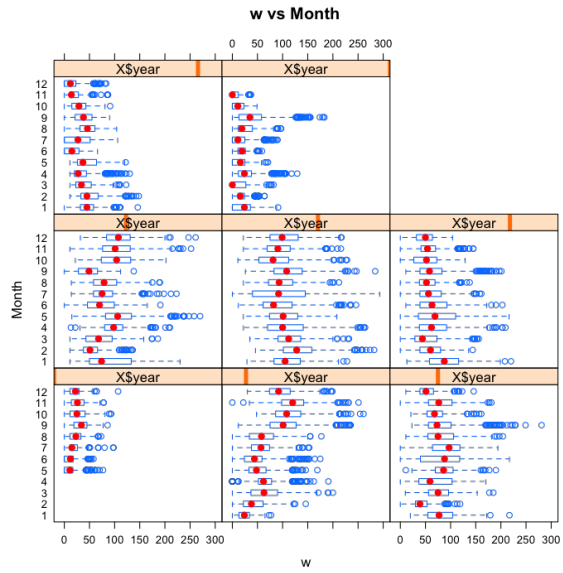


Figure 4: Box plots of raw Wolf number (w) by month and year.

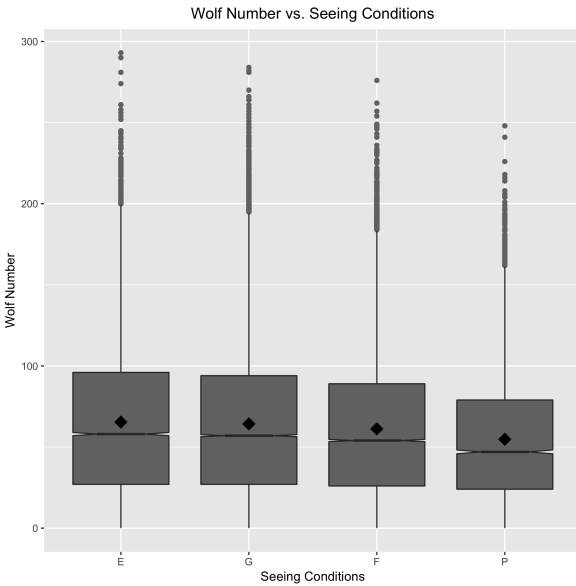


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

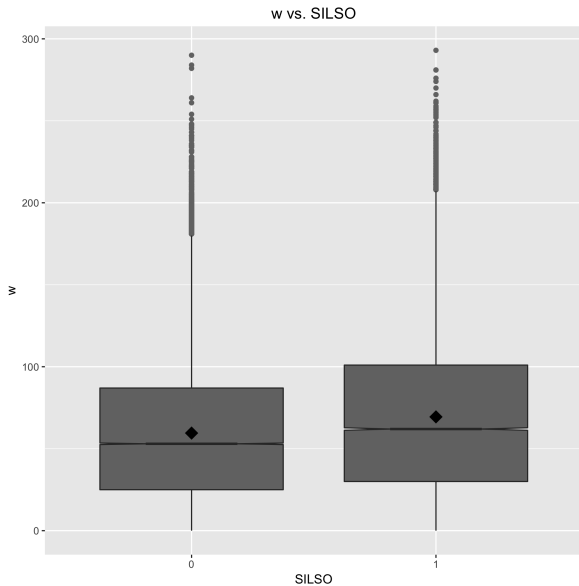


Figure 6: Box plots of raw Wolf number (w) by organization.

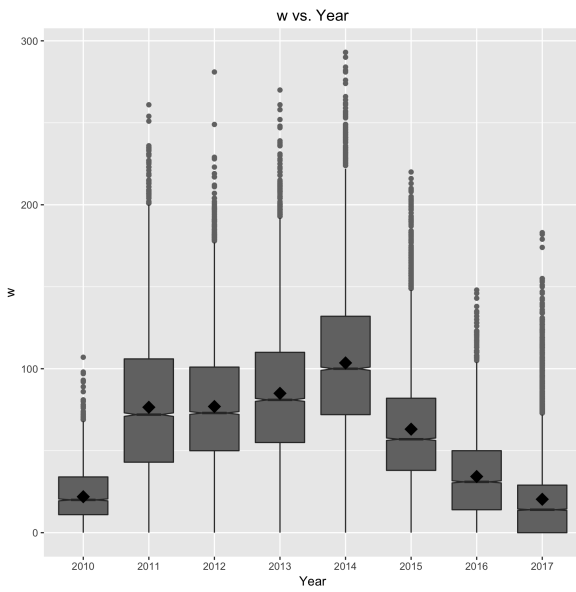


Figure 7: Box plots of raw Wolf number (w) by year.