Monthly Report (00)

201710 Data Set

Tuesday 14th November, 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.

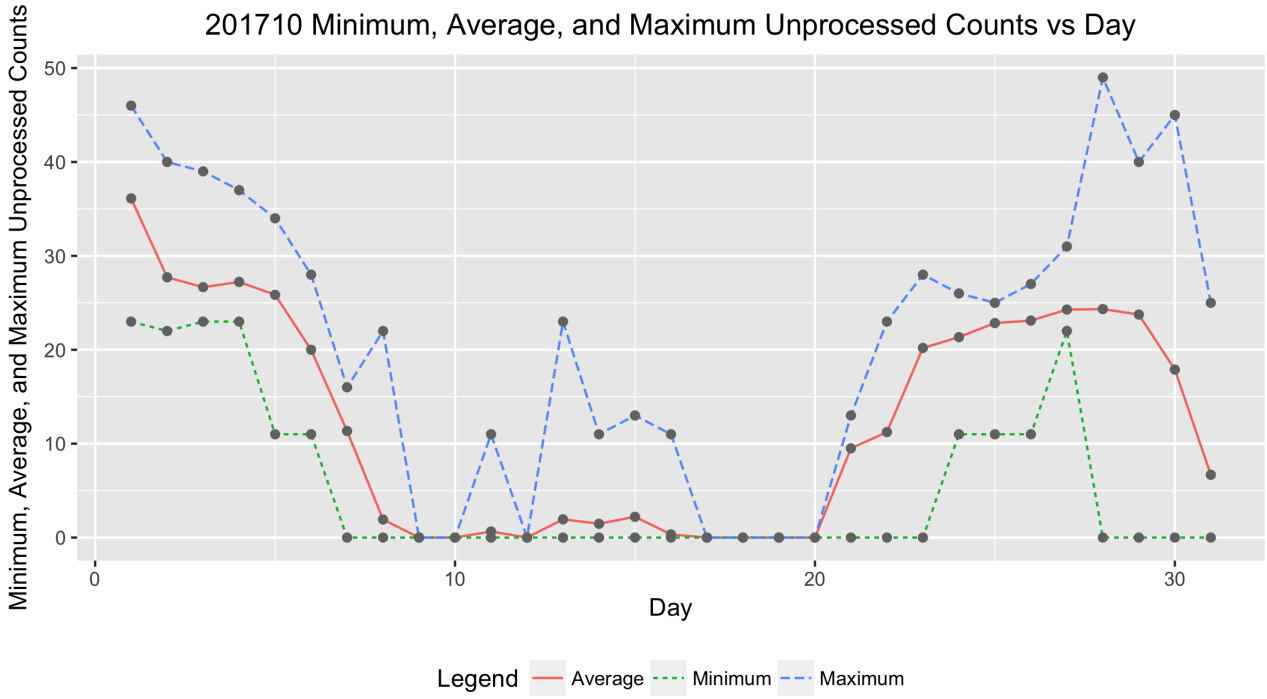

Figure 1: Raw average sunspot count by day of the month.

Table 1: 201710 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	33.0000	23.0000	36.1212	46.0000
2.0000	32.0000	22.0000	27.7188	40.0000
3.0000	39.0000	23.0000	26.6667	39.0000
4.0000	40.0000	23.0000	27.2250	37.0000
5.0000	34.0000	11.0000	25.8529	34.0000
6.0000	33.0000	11.0000	20.0000	28.0000
7.0000	32.0000	0.0000	11.3438	16.0000
8.0000	37.0000	0.0000	1.9189	22.0000
9.0000	32.0000	0.0000	0.0000	0.0000
10.0000	36.0000	0.0000	0.0000	0.0000
11.0000	34.0000	0.0000	0.6471	11.0000
12.0000	33.0000	0.0000	0.0000	0.0000
13.0000	35.0000	0.0000	1.9429	23.0000
14.0000	30.0000	0.0000	1.4667	11.0000
15.0000	33.0000	0.0000	2.2121	13.0000
16.0000	33.0000	0.0000	0.3333	11.0000
17.0000	33.0000	0.0000	0.0000	0.0000
18.0000	35.0000	0.0000	0.0000	0.0000
19.0000	31.0000	0.0000	0.0000	0.0000
20.0000	32.0000	0.0000	0.0000	0.0000
21.0000	26.0000	0.0000	9.5000	13.0000
22.0000	34.0000	0.0000	11.2353	23.0000
23.0000	31.0000	0.0000	20.1935	28.0000
24.0000	29.0000	11.0000	21.3448	26.0000
25.0000	32.0000	11.0000	22.8438	25.0000
26.0000	28.0000	11.0000	23.1071	27.0000
27.0000	32.0000	22.0000	24.2812	31.0000
28.0000	33.0000	0.0000	24.3333	49.0000
29.0000	28.0000	0.0000	23.7500	40.0000
30.0000	27.0000	0.0000	17.8889	45.0000
31.0000	35.0000	0.0000	6.6857	25.0000

3 Error Tables

Data are for the month of October 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 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	silso
2010.05	23.5975	23.0591	24.1358	8.4000	8.7000
2010.06	18.2896	17.8015	18.7776	11.0000	13.6000
2010.07	20.1891	19.7431	20.6352	15.2000	16.1000
2010.08	20.6317	20.1300	21.1333	18.3000	19.6000
2010.09	24.8270	24.2949	25.3591	22.8000	25.2000
2010.10	22.0890	21.6134	22.5647	21.0000	23.5000
2010.11	23.4558	22.9278	23.9839	20.9000	21.6000
2010.12	22.4407	21.7956	23.0858	13.9000	14.5000
2011.01	75.5389	73.8409	77.2370	17.7000	18.7000
2011.02	65.9521	64.4724	67.4318	29.1000	29.6000
2011.03	70.7451	69.2850	72.2053	48.0000	55.8000
2011.04	77.9328	76.2698	79.5958	47.3000	54.4000
2011.05	79.2753	77.7039	80.8466	37.3000	41.5000
2011.06	64.7260	63.3709	66.0812	35.2000	37.0000
2011.07	69.6455	68.1154	71.1755	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	74.5783	73.1204	76.0361	42.4000	50.5000
2011.09	86.6486	85.5108	87.7864	73.8000	78.0000
2011.10	76.5511	75.2470	77.8552	78.9000	88.0000
2011.11	79.9483	78.2371	81.6594	84.6000	96.7000
2011.12	74.3696	72.7309	76.0084	65.8000	73.0000
2012.01	78.0523	76.5066	79.5979	55.8000	58.2000
2012.02	65.9204	64.5227	67.3182	29.2000	33.1000
2012.03	73.5111	72.2013	74.8209	53.1000	64.1000
2012.04	78.1119	75.8049	80.4189	51.4000	55.2000
2012.05	83.5301	82.0787	84.9815	61.8000	69.0000
2012.06	68.0260	66.8308	69.2213	59.7000	64.5000
2012.07	74.1433	72.9003	75.3862	64.2000	51.3000
2012.08	75.6318	74.3721	76.8914	57.7000	63.1000
2012.09	88.1246	86.6188	89.6304	57.7000	61.5000
2012.10	79.2993	77.8029	80.7958	48.3000	53.3000
2012.11	83.9263	82.2435	85.6091	56.7000	61.4000
2012.12	75.8109	74.2199	77.4018	37.4000	40.8000
2013.01	87.9048	86.2686	89.5411	63.8000	62.9000
2013.02	76.0316	74.5542	77.5090	37.8000	38.0000
2013.03	80.7889	79.2743	82.3035	50.6000	57.9000
2013.04	90.8891	89.3623	92.4159	70.6000	72.4000
2013.05	91.1783	89.6122	92.7443	77.4000	78.7000
2013.06	74.8540	73.5322	76.1759	51.0000	52.5000
2013.07	79.6543	78.4183	80.8903	57.0000	57.0000
2013.08	83.3932	82.0926	84.6939	60.0000	66.0000
2013.09	96.1814	94.5200	97.8428	34.6000	36.9000
2013.10	85.0375	83.5308	86.5443	74.5000	85.6000
2013.11	90.1619	88.2914	92.0323	73.9000	77.6000
2013.12	83.4419	81.7653	85.1185	77.8000	90.3000
2014.01	104.4732	102.2811	106.6653	77.4000	82.0000
2014.02	90.2800	88.5680	91.9919	93.9000	102.8000
2014.03	99.3484	97.6767	101.0202	80.9000	92.2000
2014.04	110.4953	108.6244	112.3662	76.9000	84.7000
2014.05	110.3859	108.6220	112.1498	72.3000	75.2000
2014.06	90.6778	89.2114	92.1442	67.2000	71.0000
2014.07	97.7547	96.1633	99.3462	72.5000	72.5000
2014.08	101.8117	100.2864	103.3370	71.2000	74.7000
2014.09	118.6626	116.7503	120.5749	83.2000	87.6000
2014.10	104.7615	103.0155	106.5075	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	111.5183	109.4261	113.6106	65.8000	71.1000
2014.12	100.9722	98.7843	103.1601	75.8000	78.0000
2015.01	63.7357	62.5140	64.9575	65.9000	67.0000
2015.02	55.0514	53.7701	56.3327	42.4000	44.8000
2015.03	59.4003	58.3140	60.4866	38.0000	38.4000
2015.04	66.8376	65.6550	68.0201	49.0000	54.4000
2015.05	66.3782	65.3137	67.4428	56.3000	58.8000
2015.06	55.0071	54.0723	55.9420	50.2000	68.3000
2015.07	58.1412	57.1501	59.1324	47.9000	65.8000
2015.08	62.0755	61.0619	63.0891	39.5000	57.2000
2015.09	72.0540	70.8837	73.2244	49.2000	72.1000
2015.10	63.5388	62.4586	64.6190	39.3000	48.3000
2015.11	68.1600	67.2408	69.0793	39.6000	55.9000
2015.12	61.2489	60.0226	62.4753	36.4000	44.8000
2016.01	35.5435	34.9135	36.1735	33.7000	43.3000
2016.02	30.0910	29.4963	30.6857	38.3000	46.8000
2016.03	32.3347	31.7412	32.9281	30.5000	38.9000
2016.04	35.7883	35.1549	36.4217	26.6000	30.9000
2016.05	36.5395	35.9174	37.1616	33.7000	48.4000
2016.06	29.8889	29.4170	30.3608	13.1000	19.5000
2016.07	32.2387	31.7421	32.7352	21.2000	27.5000
2016.08	34.0837	33.5247	34.6427	33.0000	47.9000
2016.09	39.2757	38.6253	39.9260	27.7000	37.1000
2016.10	34.8980	34.2957	35.5003	22.7000	31.7000
2016.11	37.0252	36.3894	37.6609	14.0000	22.2000
2016.12	33.5734	32.9024	34.2443	11.1000	20.0000
2017.01	22.7004	22.2695	23.1314	18.4000	26.2000
2017.02	19.0351	18.6626	19.4077	14.4000	20.6000
2017.03	20.9200	20.5511	21.2888	11.3000	15.5000
2017.04	23.6050	23.2228	23.9871	21.6000	33.2000
2017.05	23.4885	23.1078	23.8692	12.5000	18.1000
2017.06	19.1403	18.8487	19.4319	15.5000	19.3000
2017.07	20.8080	20.5073	21.1087	11.5000	16.3000
2017.08	21.8156	21.4767	22.1545	22.8000	35.7000
2017.09	25.6413	25.2163	26.0664	34.6000	42.9000
2017.10	22.1499	21.7806	22.5192	10.5000	11.0000

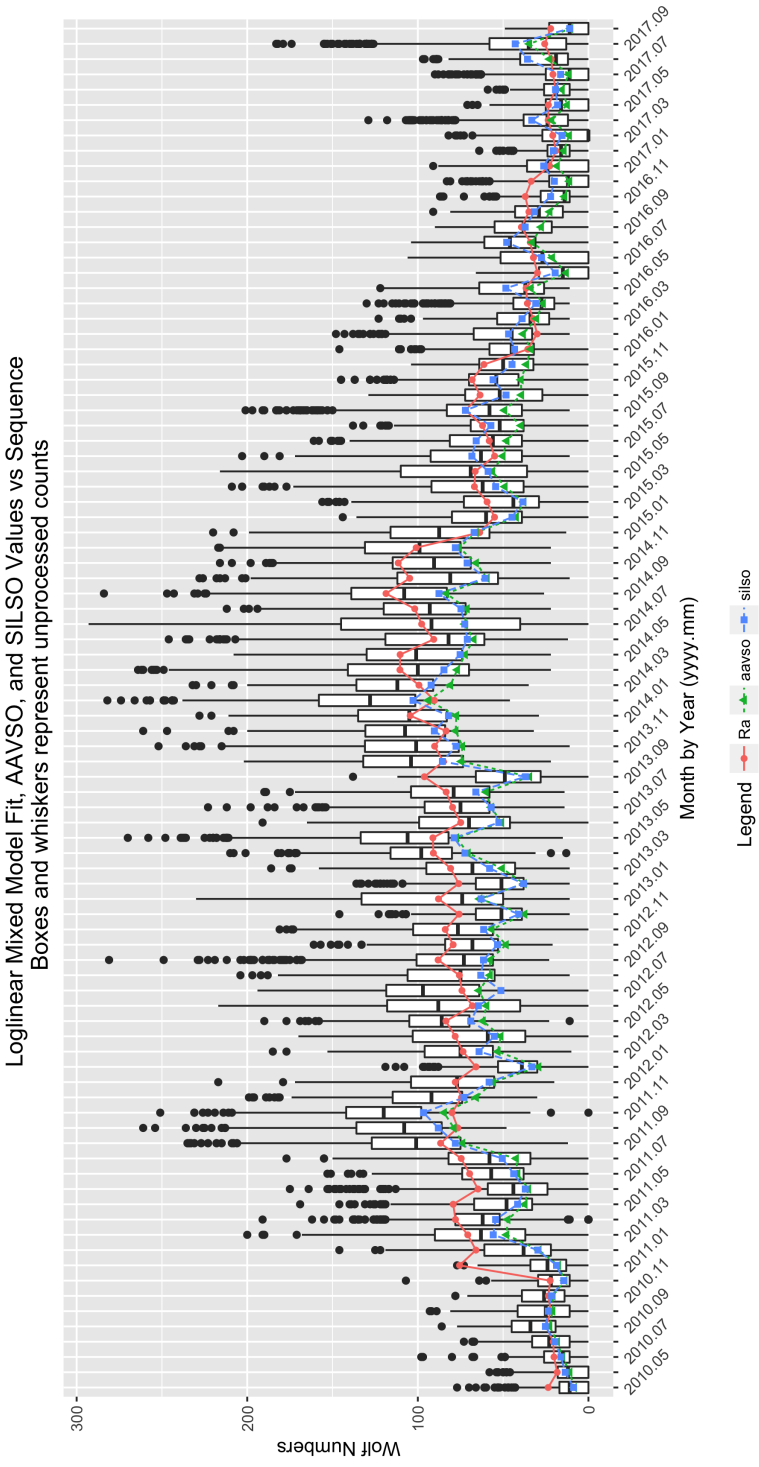

Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201710 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.2555	0.0336	96.8874	0.0000
seeF	-0.1913	0.0072	-26.7302	0.0000
seeG	-0.1026	0.0062	-16.5118	0.0000
seeP	-0.2950	0.0105	-28.0819	0.0000
silso1	-0.0191	0.0337	-0.5659	0.5715
year2011	1.2005	0.0156	77.0144	0.0000
year2012	1.2241	0.0155	78.7950	0.0000
year2013	1.3191	0.0155	85.1579	0.0000
year2014	1.5081	0.0154	98.0066	0.0000
year2015	1.0098	0.0158	63.9574	0.0000
year2016	0.4048	0.0168	24.1112	0.0000
year2017	-0.0422	0.0181	-2.3243	0.0201
mon2	-0.1552	0.0120	-12.9580	0.0000
mon3	-0.0787	0.0110	-7.1242	0.0000
mon4	0.0307	0.0110	2.7822	0.0054
mon5	0.0336	0.0105	3.2055	0.0013
mon6	-0.1730	0.0110	-15.6913	0.0000
mon7	-0.1030	0.0106	-9.6838	0.0000
mon8	-0.0509	0.0104	-4.8848	0.0000
mon9	0.1022	0.0100	10.1996	0.0000
mon10	-0.0167	0.0107	-1.5599	0.1188
mon11	0.0502	0.0110	4.5523	0.0000
mon12	-0.0390	0.0117	-3.3259	0.0009

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: 201710 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2676	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.0
CHAG : 2458	1st Qu.:2456099	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.0
BRAB : 2431	Median :2456752	Median :2014	Median : 7.000	Median :16.0
BROB : 2163	Mean :2456476	Mean :2014	Mean : 6.698	Mean :15.7
KNJS : 2073	3rd Qu.:2457427	3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.0
HOWR : 2042	Max. :2458058	Max. :2017	Max. :12.000	Max. :31.0
(Other):50425				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:12014	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24894	Min. :0.0000
F:19597	1st Qu.: 2.000	1st Qu.: 7.00	1st Qu.: 27.00	3000F : 9764	1st Qu.:0.0000
G:27463	Median : 4.000	Median : 18.00	Median : 57.00	2500E : 7766	Median :0.0000
P: 5194	Mean : 3.922	Mean : 24.37	Mean : 63.59	: 5592	Mean :0.3274
	3rd Qu.: 6.000	3rd Qu.: 36.00	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	

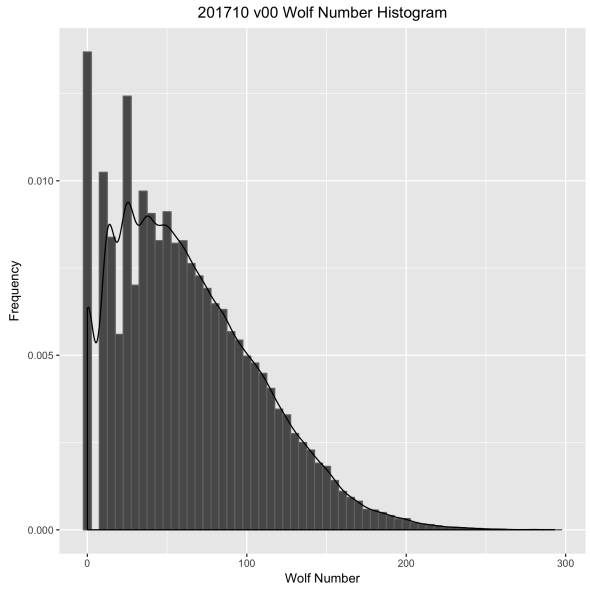

Figure 3: Box plots of raw Wolf number (w) by observer rank.

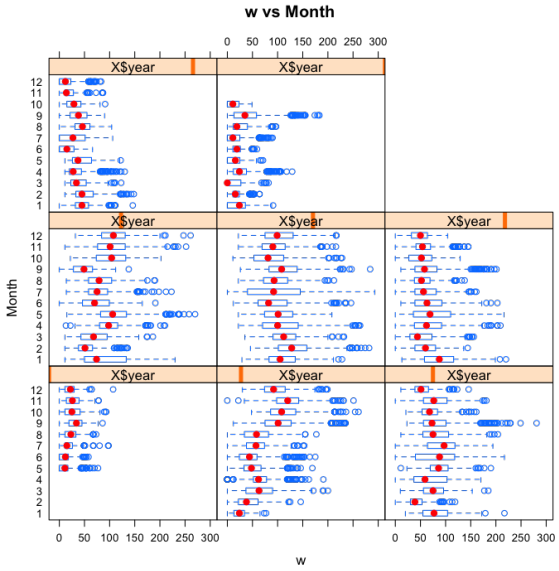

Figure 4: Box plots of raw Wolf number (w) by month and year.

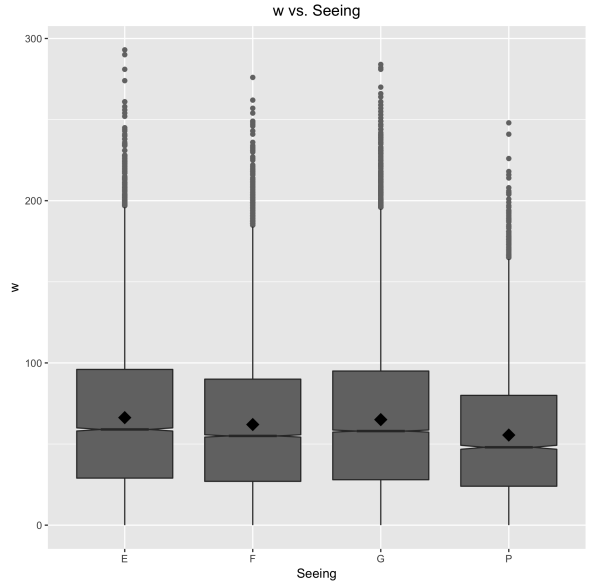

Figure 5: Box plots of raw Wolf number (w) by seeing condition.

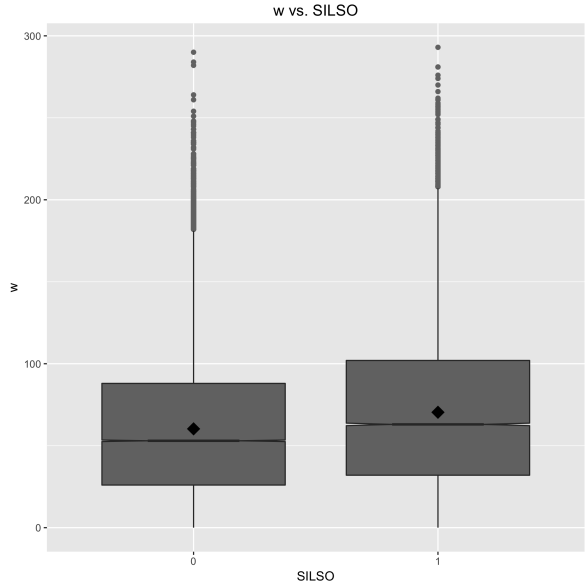

Figure 6: Box plots of raw Wolf number (w) by organization.

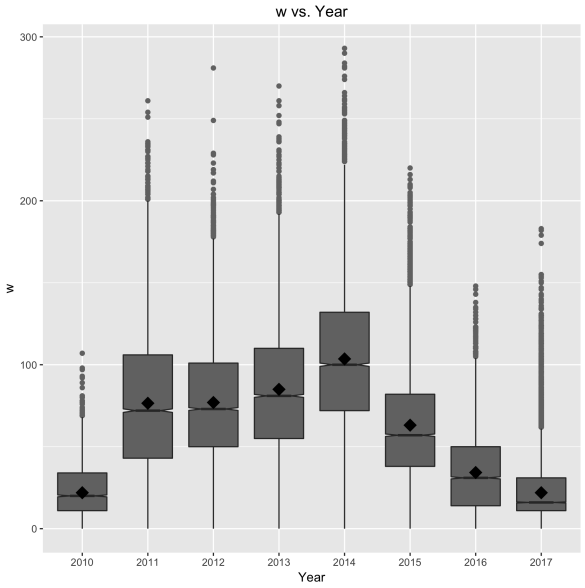

Figure 7: Box plots of raw Wolf number (w) by year.