

Monthly Report (00)

2017.04 Data Set

Saturday 13th May, 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 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.

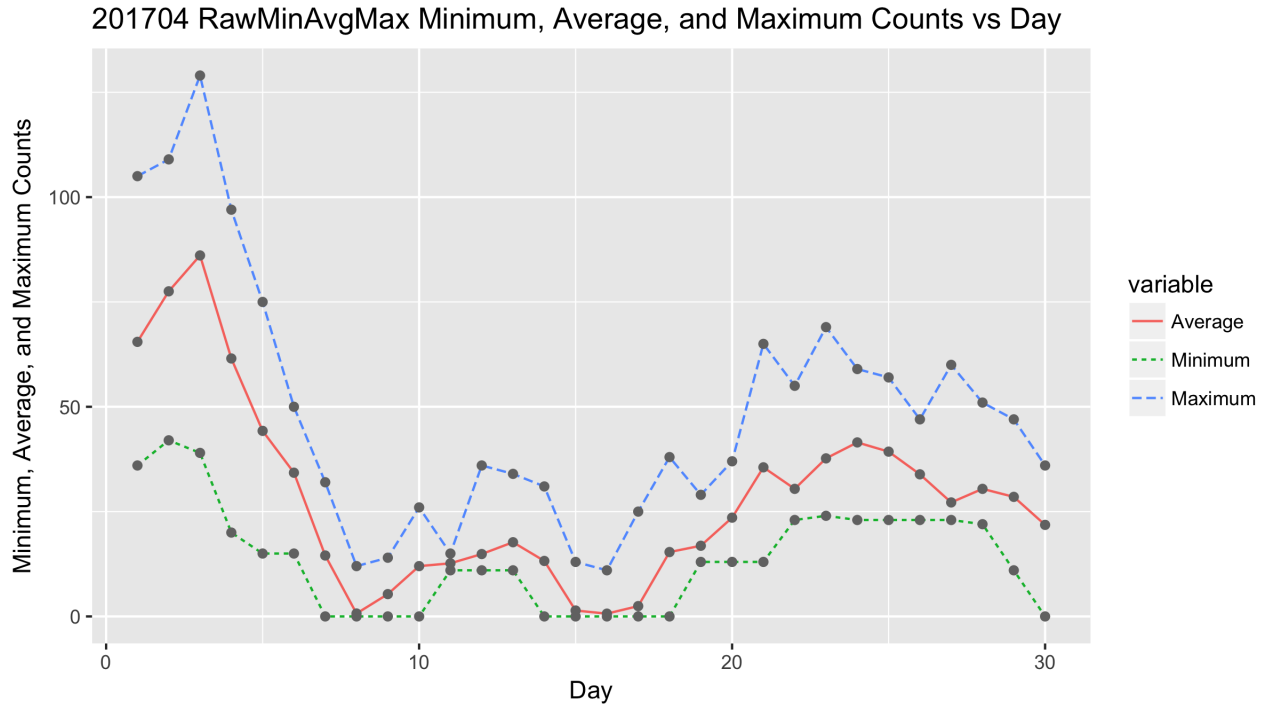


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201704 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	29.0000	36.0000	65.4828	105.0000
2.0000	33.0000	42.0000	77.5333	109.0000
3.0000	35.0000	39.0000	86.0909	129.0000
4.0000	29.0000	20.0000	61.5172	97.0000
5.0000	33.0000	15.0000	44.2500	75.0000
6.0000	31.0000	15.0000	34.2759	50.0000
7.0000	31.0000	0.0000	14.5333	32.0000
8.0000	36.0000	0.0000	0.6970	12.0000
9.0000	41.0000	0.0000	5.3243	14.0000
10.0000	36.0000	0.0000	12.0000	26.0000
11.0000	31.0000	11.0000	12.6667	15.0000
12.0000	35.0000	11.0000	14.8788	36.0000
13.0000	39.0000	11.0000	17.6944	34.0000
14.0000	39.0000	0.0000	13.2353	31.0000
15.0000	36.0000	0.0000	1.4118	13.0000
16.0000	35.0000	0.0000	0.6667	11.0000
17.0000	34.0000	0.0000	2.4643	25.0000
18.0000	34.0000	0.0000	15.3548	38.0000
19.0000	30.0000	13.0000	16.8571	29.0000
20.0000	39.0000	13.0000	23.5556	37.0000
21.0000	26.0000	13.0000	35.5600	65.0000
22.0000	35.0000	23.0000	30.4242	55.0000
23.0000	42.0000	24.0000	37.6944	69.0000
24.0000	39.0000	23.0000	41.5000	59.0000
25.0000	34.0000	23.0000	39.3030	57.0000
26.0000	27.0000	23.0000	33.8889	47.0000
27.0000	34.0000	23.0000	27.1875	60.0000
28.0000	38.0000	22.0000	30.4242	51.0000
29.0000	34.0000	11.0000	28.5152	47.0000
30.0000	32.0000	0.0000	21.8333	36.0000

3 Error Tables

Data are for the month of April 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	23.8930	23.3448	24.4411	8.4000	8.7000
2010.06	18.3096	17.8201	18.7990	11.0000	13.6000
2010.07	20.5922	20.1361	21.0483	15.2000	16.1000
2010.08	20.2963	19.8012	20.7914	18.3000	19.6000
2010.09	23.8855	23.3715	24.3994	22.8000	25.2000
2010.10	22.7071	22.2168	23.1973	21.0000	23.5000
2010.11	23.4270	22.8975	23.9564	20.9000	21.6000
2010.12	22.4176	21.7698	23.0653	13.9000	14.5000
2011.01	75.9605	74.2487	77.6722	17.7000	18.7000
2011.02	66.3078	64.8105	67.8051	29.1000	29.6000
2011.03	71.1217	69.6503	72.5930	48.0000	55.8000
2011.04	78.3579	76.6800	80.0359	47.3000	54.4000
2011.05	80.6896	79.0768	82.3024	37.3000	41.5000
2011.06	65.1409	63.7667	66.5151	35.2000	37.0000
2011.07	71.3882	69.8138	72.9626	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	73.7332	72.2850	75.1813	42.4000	50.5000
2011.09	83.9217	82.8148	85.0287	73.8000	78.0000
2011.10	79.1553	77.7918	80.5188	78.9000	88.0000
2011.11	80.3864	78.6580	82.1147	84.6000	96.7000
2011.12	74.7358	73.0876	76.3840	65.8000	73.0000
2012.01	78.0533	76.5011	79.6056	55.8000	58.2000
2012.02	65.9389	64.5319	67.3458	29.2000	33.1000
2012.03	73.5728	72.2534	74.8922	53.1000	64.1000
2012.04	78.0765	75.7654	80.3876	51.4000	55.2000
2012.05	84.6568	83.1749	86.1387	61.8000	69.0000
2012.06	68.1608	66.9490	69.3726	59.7000	64.5000
2012.07	75.6719	74.3932	76.9506	64.2000	51.3000
2012.08	74.4453	73.1957	75.6949	57.7000	63.1000
2012.09	84.7644	83.3109	86.2178	57.7000	61.5000
2012.10	81.5291	79.9860	83.0721	48.3000	53.3000
2012.11	83.8481	82.1645	85.5317	56.7000	61.4000
2012.12	75.7105	74.1143	77.3066	37.4000	40.8000
2013.01	87.9446	86.2976	89.5915	63.8000	62.9000
2013.02	76.0396	74.5544	77.5249	37.8000	38.0000
2013.03	80.7665	79.2428	82.2902	50.6000	57.9000
2013.04	90.9134	89.3802	92.4466	70.6000	72.4000
2013.05	92.4017	90.8052	93.9982	77.4000	78.7000
2013.06	75.0191	73.6809	76.3573	51.0000	52.5000
2013.07	81.2333	79.9658	82.5007	57.0000	57.0000
2013.08	82.0501	80.7658	83.3344	60.0000	66.0000
2013.09	92.5448	90.9419	94.1477	34.6000	36.9000
2013.10	87.4534	85.8995	89.0074	74.5000	85.6000
2013.11	90.0551	88.1799	91.9302	73.9000	77.6000
2013.12	83.4049	81.7262	85.0836	77.8000	90.3000
2014.01	104.4299	102.2315	106.6284	77.4000	82.0000
2014.02	90.2582	88.5410	91.9755	93.9000	102.8000
2014.03	99.3272	97.6472	101.0073	80.9000	92.2000
2014.04	110.4794	108.5991	112.3597	76.9000	84.7000
2014.05	111.8429	110.0490	113.6369	72.3000	75.2000
2014.06	90.7882	89.3130	92.2634	67.2000	71.0000
2014.07	99.6772	98.0483	101.3060	72.5000	72.5000
2014.08	100.1798	98.6709	101.6888	71.2000	74.7000
2014.09	114.1758	112.3298	116.0219	83.2000	87.6000
2014.10	107.6685	105.8636	109.4733	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	111.4126	109.3080	113.5172	65.8000	71.1000
2014.12	100.8483	98.6573	103.0392	75.8000	78.0000
2015.01	63.7198	62.4882	64.9515	65.9000	67.0000
2015.02	55.0612	53.7730	56.3494	42.4000	44.8000
2015.03	59.3875	58.2946	60.4804	38.0000	38.4000
2015.04	66.8752	65.6857	68.0646	49.0000	54.4000
2015.05	67.3059	66.2226	68.3893	56.3000	58.8000
2015.06	55.1081	54.1662	56.0501	50.2000	68.3000
2015.07	59.3062	58.2898	60.3225	47.9000	65.8000
2015.08	61.0651	60.0609	62.0693	39.5000	57.2000
2015.09	69.3058	68.1742	70.4375	49.2000	72.1000
2015.10	65.3270	64.2103	66.4438	39.3000	48.3000
2015.11	68.1047	67.1807	69.0287	39.6000	55.9000
2015.12	61.1636	59.9262	62.4009	36.4000	44.8000
2016.01	35.5576	34.9251	36.1901	33.7000	43.3000
2016.02	30.0833	29.4852	30.6814	38.3000	46.8000
2016.03	32.3342	31.7386	32.9299	30.5000	38.9000
2016.04	35.7852	35.1476	36.4227	26.6000	30.9000
2016.05	37.0111	36.3765	37.6458	33.7000	48.4000
2016.06	29.9275	29.4511	30.4039	13.1000	19.5000
2016.07	32.8967	32.3868	33.4067	21.2000	27.5000
2016.08	33.5222	32.9682	34.0762	33.0000	47.9000
2016.09	37.8134	37.1839	38.4428	27.7000	37.1000
2016.10	35.8801	35.2573	36.5029	22.7000	31.7000
2016.11	36.9915	36.3534	37.6296	14.0000	22.2000
2016.12	33.5317	32.8576	34.2058	11.1000	20.0000
2017.01	22.2880	21.8627	22.7132	18.4000	26.2000
2017.02	18.6878	18.3211	19.0544	14.4000	20.6000
2017.03	20.5338	20.1696	20.8980	11.3000	15.5000
2017.04	23.1852	22.8078	23.5626	21.6000	33.2000

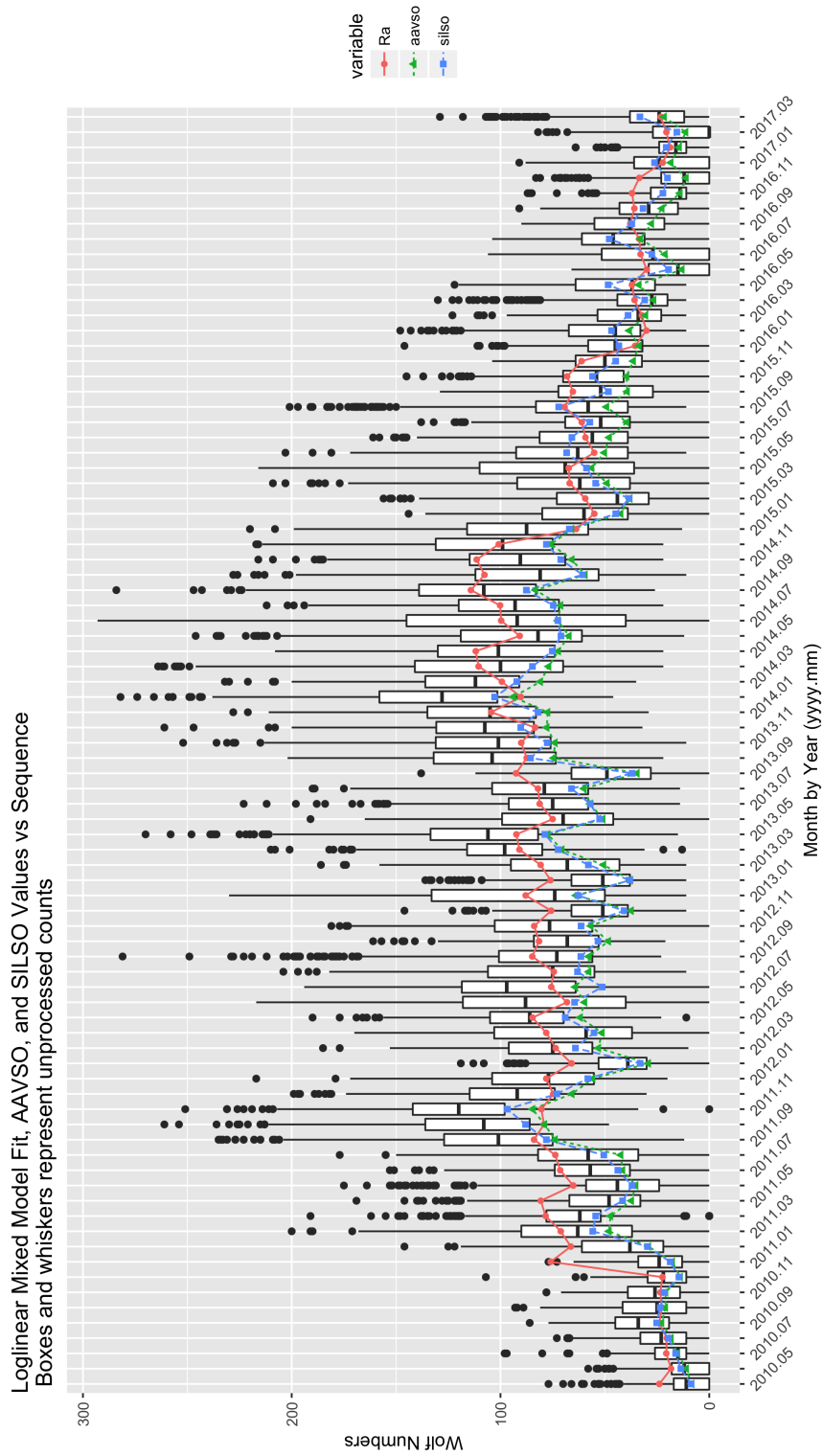


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201704 Parameter Estimates

	Estimate	Std. Error	t-value	$Pr(> t)$
(Intercept)	3.2140	0.0441	72.9000	0.0000
seeF	-0.1892	0.0072	-26.2029	0.0000
seeG	-0.1022	0.0063	-16.2750	0.0000
seeP	-0.2942	0.0106	-27.7869	0.0000
r1000B	-0.0553	0.0827	-0.6685	0.5038
r1500C	0.0393	0.1266	0.3101	0.7565
r2000D	0.0787	0.1545	0.5096	0.6103
r2500E	0.0005	0.1048	0.0048	0.9962
r3000F	0.0760	0.1021	0.7448	0.4564
r3500G	0.1243	0.1529	0.8133	0.4160
r5000H	-0.1010	0.2114	-0.4776	0.6329
silsoy	0.1236	0.0735	1.6816	0.0926
year2011	1.2059	0.0155	77.9793	0.0000
year2012	1.2235	0.0154	79.3818	0.0000
year2013	1.3187	0.0154	85.8042	0.0000
year2014	1.5072	0.0153	98.7283	0.0000
year2015	1.0092	0.0157	64.4334	0.0000
year2016	0.4042	0.0167	24.2672	0.0000
year2017	-0.0613	0.0239	-2.5678	0.0102
mon2	-0.1553	0.0119	-13.0769	0.0000
mon3	-0.0789	0.0110	-7.2027	0.0000
mon4	0.0307	0.0109	2.8025	0.0051
mon5	0.0464	0.0105	4.4163	0.0000
mon6	-0.1713	0.0111	-15.3798	0.0000
mon7	-0.0831	0.0107	-7.7835	0.0000
mon8	-0.0672	0.0105	-6.3936	0.0000
mon9	0.0635	0.0101	6.2636	0.0000
mon10	0.0109	0.0107	1.0147	0.3103
mon11	0.0491	0.0110	4.4728	0.0000
mon12	-0.0402	0.0117	-3.4506	0.0006

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: 201704 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2492	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2283	1st Qu.:2456021	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2277	Median :2456594	Median :2013	Median : 7.000	Median :16.00
BROB : 2005	Mean :2456313	Mean :2013	Mean : 6.602	Mean :15.72
KNJS : 1896	3rd Qu.:2457227	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 1885	Max. :2457874	Max. :2017	Max. :12.000	Max. :31.00
(Other):45087				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:10521	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24585	n:38875
F:17871	1st Qu.: 2.000	1st Qu.: 9.00	1st Qu.: 34.00	3000F : 9623	y:19050
G:24729	Median : 4.000	Median : 20.00	Median : 62.00	2500E : 7659	
P: 4804	Mean : 4.201	Mean : 26.12	Mean : 68.12	3500G : 4560	
	3rd Qu.: 6.000	3rd Qu.: 38.00	3rd Qu.: 96.00	1000B : 4181	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3043	
				(Other): 4274	

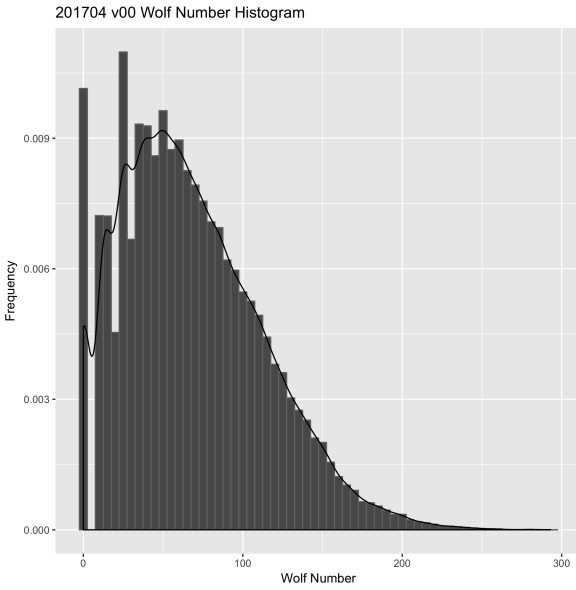


Figure 3: Box plots of raw Wolf number (w) by observer rank.

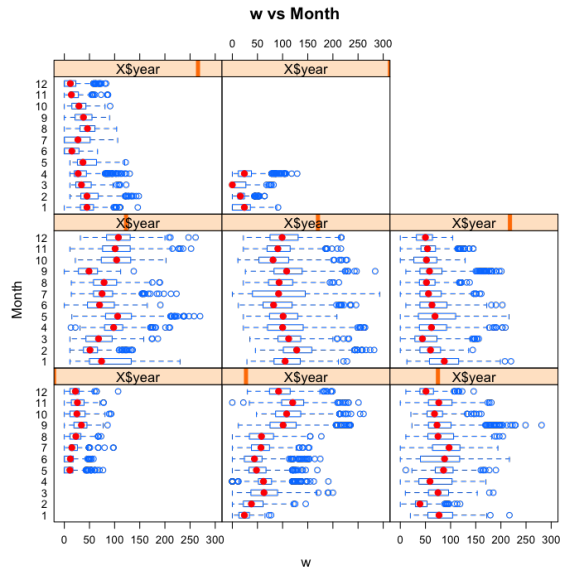


Figure 4: Box plots of raw Wolf number (w) by month and year.

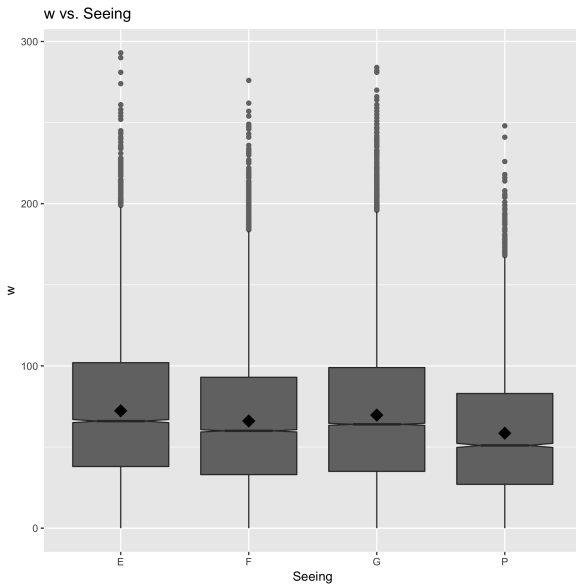


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

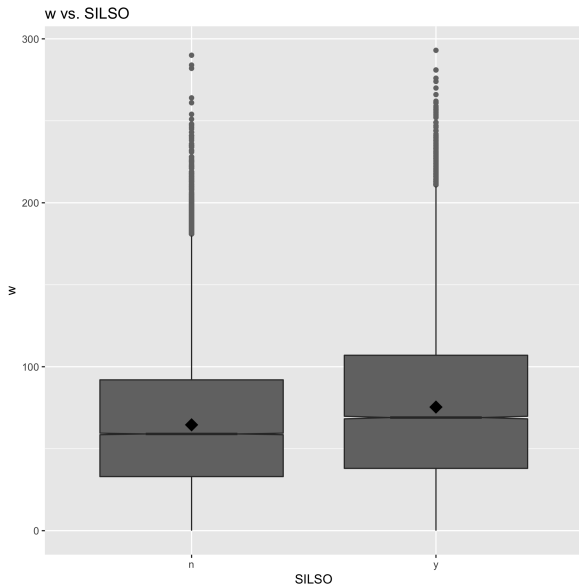


Figure 6: Box plots of raw Wolf number (w) by organization.

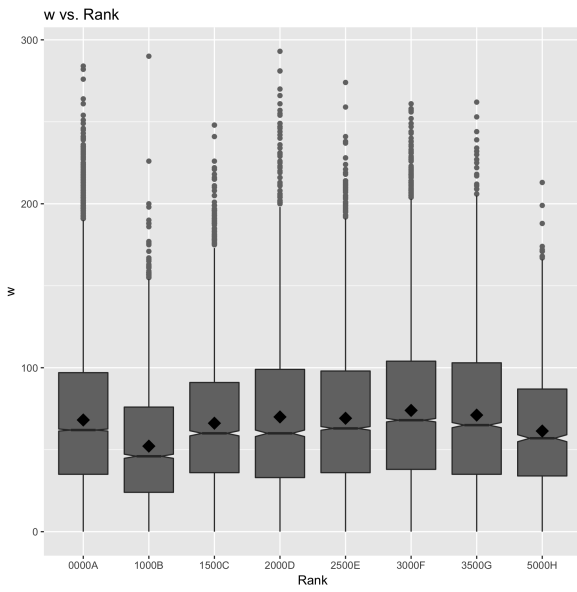


Figure 7: Box plots of raw Wolf number (w) by observer rank.

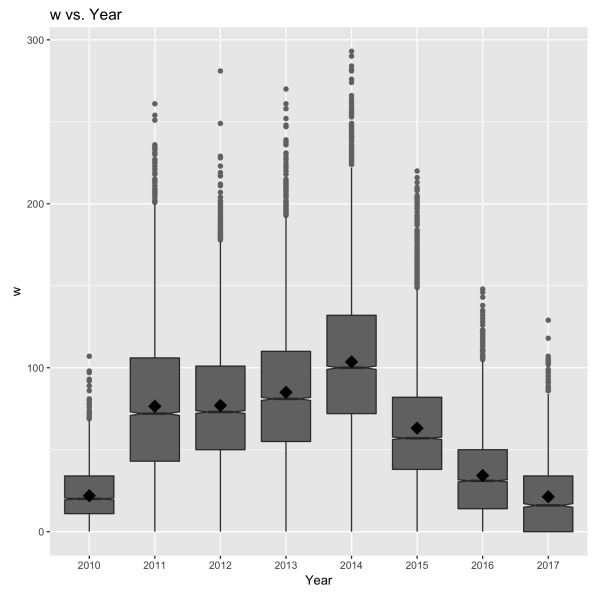


Figure 8: Box plots of raw Wolf number (w) by year.