

Monthly Report (00)

2017.03 Data Set

Wednesday 19th April, 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.

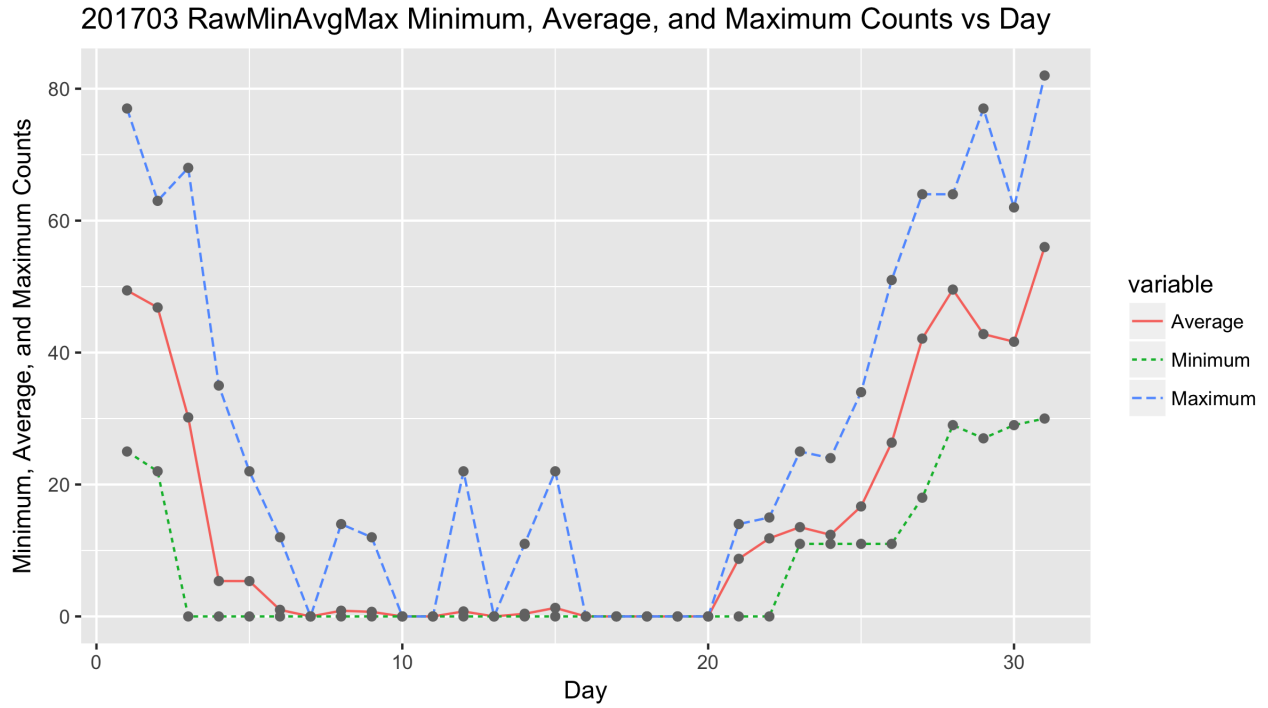


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201703 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	34.0000	25.0000	49.4194	77.0000
2.0000	34.0000	22.0000	46.8333	63.0000
3.0000	27.0000	0.0000	30.1739	68.0000
4.0000	29.0000	0.0000	5.3704	35.0000
5.0000	33.0000	0.0000	5.3571	22.0000
6.0000	25.0000	0.0000	1.0000	12.0000
7.0000	31.0000	0.0000	0.0000	0.0000
8.0000	34.0000	0.0000	0.8667	14.0000
9.0000	36.0000	0.0000	0.6970	12.0000
10.0000	29.0000	0.0000	0.0000	0.0000
11.0000	30.0000	0.0000	0.0000	0.0000
12.0000	33.0000	0.0000	0.7586	22.0000
13.0000	23.0000	0.0000	0.0000	0.0000
14.0000	28.0000	0.0000	0.4074	11.0000
15.0000	38.0000	0.0000	1.2941	22.0000
16.0000	33.0000	0.0000	0.0000	0.0000
17.0000	34.0000	0.0000	0.0000	0.0000
18.0000	28.0000	0.0000	0.0000	0.0000
19.0000	32.0000	0.0000	0.0000	0.0000
20.0000	27.0000	0.0000	0.0000	0.0000
21.0000	25.0000	0.0000	8.7273	14.0000
22.0000	30.0000	0.0000	11.8400	15.0000
23.0000	33.0000	11.0000	13.5357	25.0000
24.0000	30.0000	11.0000	12.3750	24.0000
25.0000	31.0000	11.0000	16.6786	34.0000
26.0000	30.0000	11.0000	26.3462	51.0000
27.0000	30.0000	18.0000	42.1200	64.0000
28.0000	34.0000	29.0000	49.5517	64.0000
29.0000	34.0000	27.0000	42.8000	77.0000
30.0000	29.0000	29.0000	41.6429	62.0000
31.0000	31.0000	30.0000	56.0000	82.0000

3 Error Tables

Data are for the month of March 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	23.8933	23.3452	24.4415	8.4000	8.7000
2010.06	18.3090	17.8195	18.7985	11.0000	13.6000
2010.07	20.5926	20.1364	21.0488	15.2000	16.1000
2010.08	20.2992	19.8042	20.7943	18.3000	19.6000
2010.09	23.8862	23.3722	24.4001	22.8000	25.2000
2010.10	22.7039	22.2135	23.1943	21.0000	23.5000
2010.11	23.4250	22.8954	23.9545	20.9000	21.6000
2010.12	22.4193	21.7714	23.0671	13.9000	14.5000
2011.01	76.2694	74.5516	77.9872	17.7000	18.7000
2011.02	66.5657	65.0634	68.0681	29.1000	29.6000
2011.03	71.4030	69.9271	72.8789	48.0000	55.8000
2011.04	77.5599	75.8992	79.2207	47.3000	54.4000
2011.05	80.6884	79.0759	82.3008	37.3000	41.5000
2011.06	65.1385	63.7645	66.5125	35.2000	37.0000
2011.07	71.3904	69.8165	72.9644	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	73.7325	72.2857	75.1794	42.4000	50.5000
2011.09	83.9293	82.8228	85.0359	73.8000	78.0000
2011.10	79.1619	77.7995	80.5244	78.9000	88.0000
2011.11	80.3966	78.6692	82.1240	84.6000	96.7000
2011.12	74.7452	73.0979	76.3925	65.8000	73.0000
2012.01	78.3207	76.7638	79.8776	55.8000	58.2000
2012.02	66.1463	64.7359	67.5567	29.2000	33.1000
2012.03	73.8148	72.4916	75.1380	53.1000	64.1000
2012.04	77.2278	74.9426	79.5130	51.4000	55.2000
2012.05	84.6013	83.1210	86.0815	61.8000	69.0000
2012.06	68.1153	66.9046	69.3259	59.7000	64.5000
2012.07	75.6231	74.3457	76.9005	64.2000	51.3000
2012.08	74.3988	73.1506	75.6471	57.7000	63.1000
2012.09	84.7069	83.2550	86.1588	57.7000	61.5000
2012.10	81.4692	79.9278	83.0105	48.3000	53.3000
2012.11	83.7873	82.1052	85.4695	56.7000	61.4000
2012.12	75.6573	74.0624	77.2522	37.4000	40.8000
2013.01	88.3343	86.6795	89.9891	63.8000	62.9000
2013.02	76.3488	74.8578	77.8397	37.8000	38.0000
2013.03	81.1009	79.5710	82.6307	50.6000	57.9000
2013.04	90.0035	88.4863	91.5207	70.6000	72.4000
2013.05	92.4226	90.8264	94.0189	77.4000	78.7000
2013.06	75.0343	73.6961	76.3724	51.0000	52.5000
2013.07	81.2447	79.9774	82.5120	57.0000	57.0000
2013.08	82.0668	80.7827	83.3509	60.0000	66.0000
2013.09	92.5605	90.9585	94.1625	34.6000	36.9000
2013.10	87.4781	85.9239	89.0322	74.5000	85.6000
2013.11	90.0734	88.1985	91.9482	73.9000	77.6000
2013.12	83.4225	81.7442	85.1008	77.8000	90.3000
2014.01	104.8794	102.6714	107.0874	77.4000	82.0000
2014.02	90.6304	88.9067	92.3541	93.9000	102.8000
2014.03	99.7282	98.0421	101.4142	80.9000	92.2000
2014.04	109.3621	107.5013	111.2230	76.9000	84.7000
2014.05	111.8605	110.0670	113.6540	72.3000	75.2000
2014.06	90.8061	89.3316	92.2806	67.2000	71.0000
2014.07	99.6993	98.0710	101.3276	72.5000	72.5000
2014.08	100.1970	98.6882	101.7058	71.2000	74.7000
2014.09	114.1917	112.3472	116.0363	83.2000	87.6000
2014.10	107.6857	105.8807	109.4907	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	111.4272	109.3215	113.5328	65.8000	71.1000
2014.12	100.8560	98.6673	103.0448	75.8000	78.0000
2015.01	63.9932	62.7571	65.2292	65.9000	67.0000
2015.02	55.2909	53.9981	56.5837	42.4000	44.8000
2015.03	59.6347	58.5380	60.7313	38.0000	38.4000
2015.04	66.2074	65.0302	67.3846	49.0000	54.4000
2015.05	67.3237	66.2411	68.4063	56.3000	58.8000
2015.06	55.1215	54.1801	56.0629	50.2000	68.3000
2015.07	59.3191	58.3033	60.3350	47.9000	65.8000
2015.08	61.0780	60.0741	62.0818	39.5000	57.2000
2015.09	69.3159	68.1849	70.4468	49.2000	72.1000
2015.10	65.3360	64.2197	66.4524	39.3000	48.3000
2015.11	68.1151	67.1918	69.0383	39.6000	55.9000
2015.12	61.1719	59.9343	62.4095	36.4000	44.8000
2016.01	35.7085	35.0741	36.3429	33.7000	43.3000
2016.02	30.2020	29.6020	30.8020	38.3000	46.8000
2016.03	32.4652	31.8680	33.0624	30.5000	38.9000
2016.04	35.4232	34.7927	36.0538	26.6000	30.9000
2016.05	37.0150	36.3807	37.6493	33.7000	48.4000
2016.06	29.9326	29.4564	30.4088	13.1000	19.5000
2016.07	32.9020	32.3923	33.4117	21.2000	27.5000
2016.08	33.5269	32.9731	34.0808	33.0000	47.9000
2016.09	37.8180	37.1889	38.4472	27.7000	37.1000
2016.10	35.8866	35.2640	36.5091	22.7000	31.7000
2016.11	36.9997	36.3618	37.6376	14.0000	22.2000
2016.12	33.5426	32.8684	34.2168	11.1000	20.0000
2017.01	20.3700	19.9813	20.7586	18.4000	26.2000
2017.02	17.0748	16.7402	17.4094	14.4000	20.6000
2017.03	18.7609	18.4284	19.0934	11.3000	15.5000

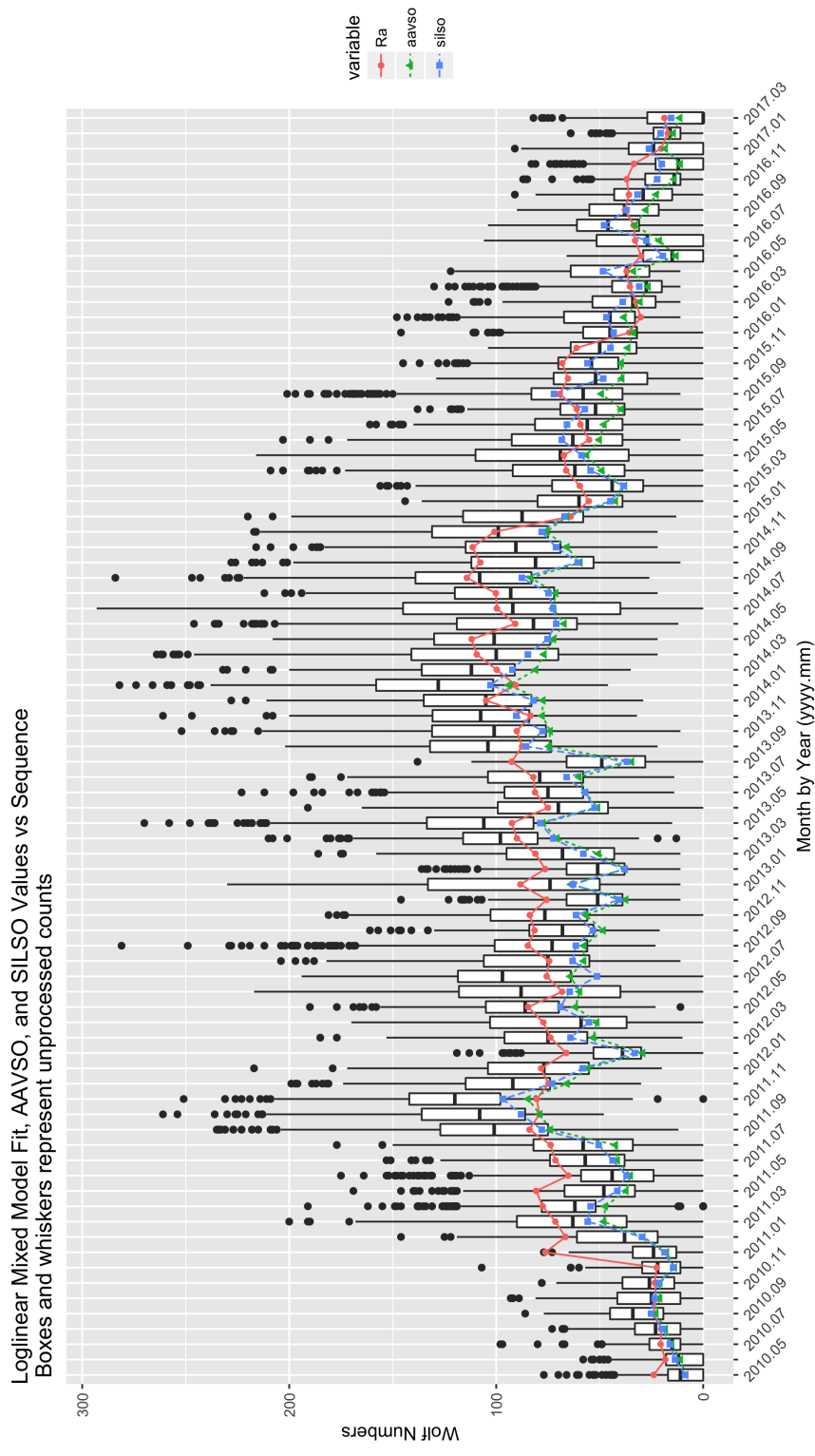


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201703 Parameter Estimates

	Estimate	Std. Error	t-value	$Pr(> t)$
(Intercept)	3.2211	0.0446	72.3039	0.0000
seeF	-0.1897	0.0072	-26.2384	0.0000
seeG	-0.1027	0.0063	-16.3405	0.0000
seeP	-0.2950	0.0106	-27.8322	0.0000
r1000B	-0.0577	0.0832	-0.6939	0.4878
r1500C	0.0367	0.1272	0.2887	0.7728
r2000D	0.0774	0.1551	0.4992	0.6176
r2500E	-0.0008	0.1053	-0.0076	0.9939
r3000F	0.0736	0.1025	0.7179	0.4728
r3500G	0.1227	0.1535	0.7993	0.4241
r5000H	-0.1050	0.2123	-0.4945	0.6210
silsoy	0.1219	0.0739	1.6502	0.0989
year2011	1.2060	0.0154	78.0963	0.0000
year2012	1.2228	0.0154	79.4446	0.0000
year2013	1.3189	0.0153	85.9349	0.0000
year2014	1.5074	0.0152	98.8742	0.0000
year2015	1.0095	0.0156	64.5396	0.0000
year2016	0.4044	0.0166	24.3124	0.0000
year2017	-0.1554	0.0278	-5.5920	0.0000
mon2	-0.1555	0.0119	-13.1128	0.0000
mon3	-0.0791	0.0109	-7.2317	0.0000
mon4	0.0162	0.0111	1.4575	0.1450
mon5	0.0423	0.0105	4.0220	0.0001
mon6	-0.1755	0.0111	-15.7619	0.0000
mon7	-0.0873	0.0107	-8.1753	0.0000
mon8	-0.0714	0.0105	-6.7910	0.0000
mon9	0.0593	0.0101	5.8540	0.0000
mon10	0.0067	0.0107	0.6273	0.5305
mon11	0.0449	0.0110	4.0959	0.0000
mon12	-0.0444	0.0117	-3.8055	0.0001

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: 201703 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2463	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2258	1st Qu.:2456014	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2248	Median :2456578	Median :2013	Median : 7.000	Median :16.00
BROB : 1980	Mean :2456294	Mean :2013	Mean : 6.633	Mean :15.72
KNJS : 1866	3rd Qu.:2457205	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 1864	Max. :2457844	Max. :2017	Max. :12.000	Max. :31.00
(Other):44545				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:10395	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24299	n:38431
F:17646	1st Qu.: 2.000	1st Qu.: 9.00	1st Qu.: 35.00	3000F : 9496	y:18793
G:24432	Median : 4.000	Median : 20.00	Median : 62.00	2500E : 7557	
P: 4751	Mean : 4.231	Mean : 26.31	Mean : 68.62	3500G : 4506	
	3rd Qu.: 6.000	3rd Qu.: 38.00	3rd Qu.: 97.00	1000B : 4128	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3032	
				(Other): 4206	

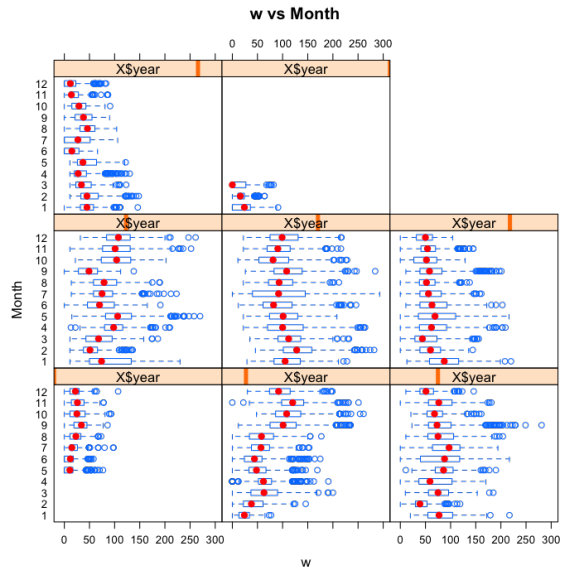
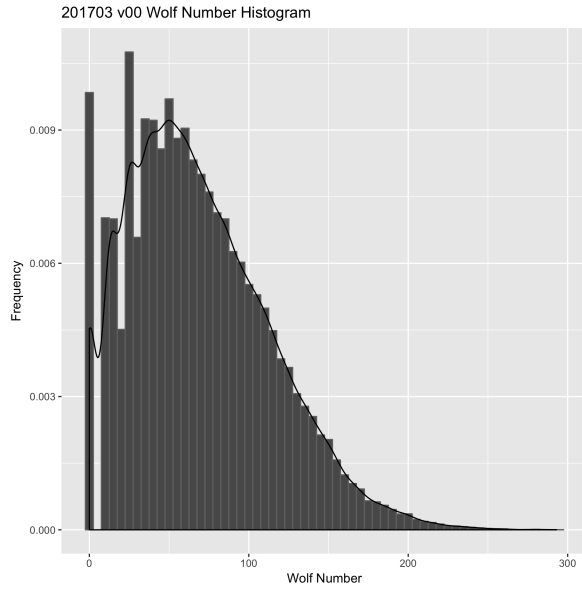


Figure 3: Box plots of raw Wolf number (w) by observer rank.

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

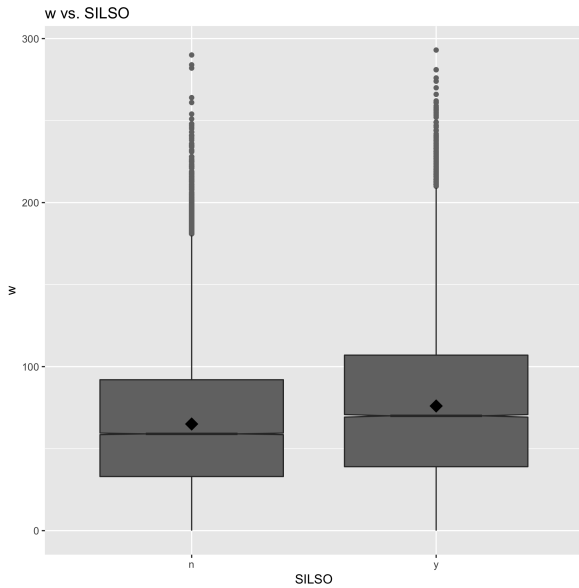
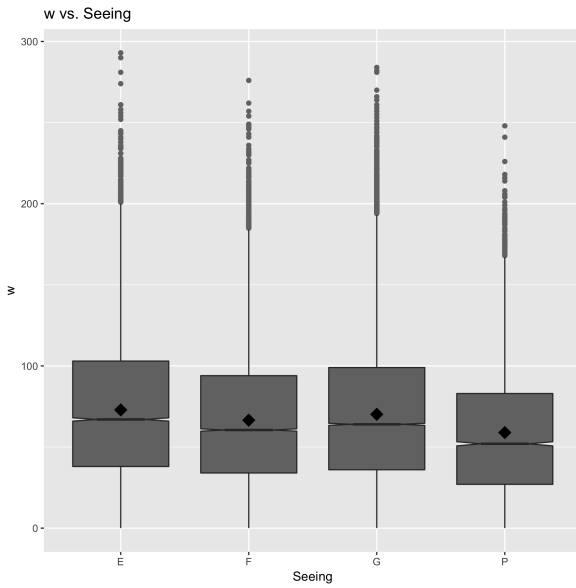


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

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

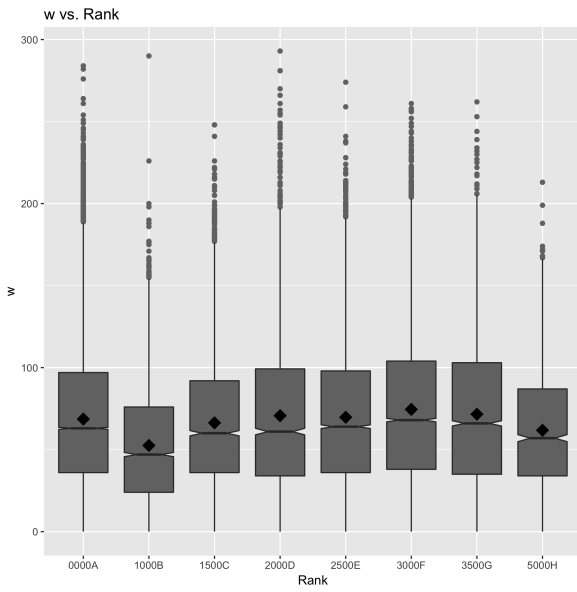


Figure 7: Box plots of raw Wolf number (w) by observer rank.

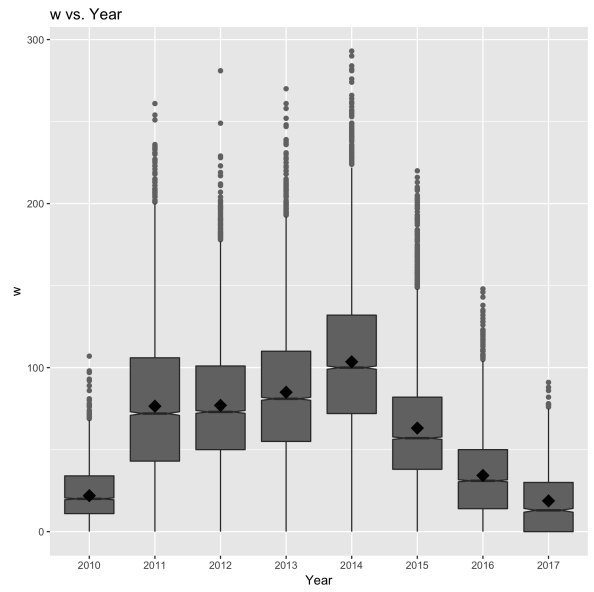


Figure 8: Box plots of raw Wolf number (w) by year.