

Monthly Report (00)

201706 Data Set

Thursday 13th July, 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.

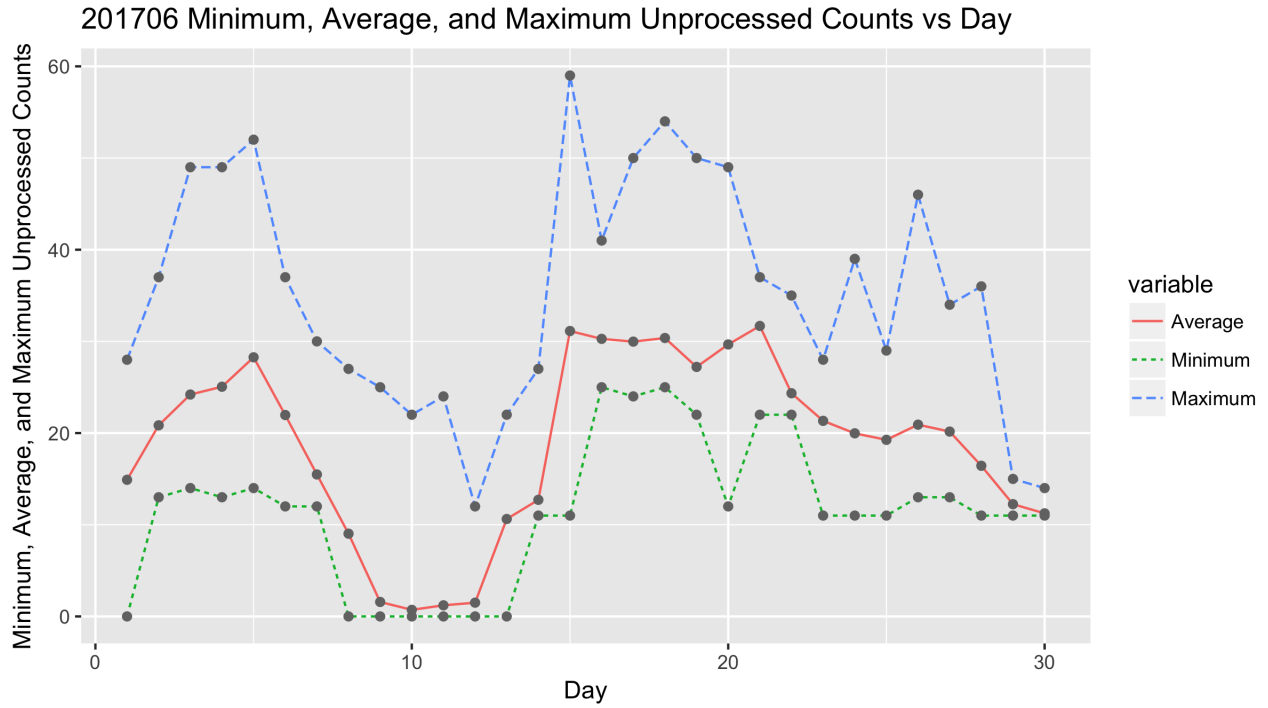


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201706 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	42.0000	0.0000	14.9048	28.0000
2.0000	39.0000	13.0000	20.8462	37.0000
3.0000	44.0000	14.0000	24.2045	49.0000
4.0000	37.0000	13.0000	25.0541	49.0000
5.0000	30.0000	14.0000	28.2667	52.0000
6.0000	33.0000	12.0000	21.9697	37.0000
7.0000	35.0000	12.0000	15.4857	30.0000
8.0000	38.0000	0.0000	9.0263	27.0000
9.0000	38.0000	0.0000	1.5789	25.0000
10.0000	46.0000	0.0000	0.7174	22.0000
11.0000	47.0000	0.0000	1.2128	24.0000
12.0000	38.0000	0.0000	1.5000	12.0000
13.0000	40.0000	0.0000	10.6250	22.0000
14.0000	39.0000	11.0000	12.7179	27.0000
15.0000	38.0000	11.0000	31.1316	59.0000
16.0000	40.0000	25.0000	30.2750	41.0000
17.0000	41.0000	24.0000	29.9756	50.0000
18.0000	38.0000	25.0000	30.3684	54.0000
19.0000	42.0000	22.0000	27.2143	50.0000
20.0000	42.0000	12.0000	29.6667	49.0000
21.0000	38.0000	22.0000	31.6842	37.0000
22.0000	32.0000	22.0000	24.3438	35.0000
23.0000	33.0000	11.0000	21.3333	28.0000
24.0000	36.0000	11.0000	19.9722	39.0000
25.0000	42.0000	11.0000	19.2619	29.0000
26.0000	49.0000	13.0000	20.9184	46.0000
27.0000	37.0000	13.0000	20.1622	34.0000
28.0000	37.0000	11.0000	16.4324	36.0000
29.0000	41.0000	11.0000	12.2439	15.0000
30.0000	37.0000	11.0000	11.2432	14.0000

3 Error Tables

Data are for the month of June 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	23.6771	23.1373	24.2169	8.4000	8.7000
2010.06	18.3859	17.8968	18.8750	11.0000	13.6000
2010.07	20.6128	20.1583	21.0673	15.2000	16.1000
2010.08	20.3170	19.8237	20.8103	18.3000	19.6000
2010.09	23.9106	23.3989	24.4223	22.8000	25.2000
2010.10	22.7334	22.2450	23.2219	21.0000	23.5000
2010.11	23.4543	22.9273	23.9813	20.9000	21.6000
2010.12	22.4281	21.7819	23.0743	13.9000	14.5000
2011.01	76.1685	74.4550	77.8819	17.7000	18.7000
2011.02	66.4790	64.9822	67.9759	29.1000	29.6000
2011.03	71.3016	69.8293	72.7738	48.0000	55.8000
2011.04	78.6047	76.9263	80.2832	47.3000	54.4000
2011.05	79.8776	78.2865	81.4686	37.3000	41.5000
2011.06	65.3251	63.9522	66.6980	35.2000	37.0000
2011.07	71.3799	69.8104	72.9494	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	73.7195	72.2756	75.1634	42.4000	50.5000
2011.09	83.9009	82.7985	85.0033	73.8000	78.0000
2011.10	79.1393	77.7815	80.4972	78.9000	88.0000
2011.11	80.3451	78.6221	82.0681	84.6000	96.7000
2011.12	74.6987	73.0532	76.3442	65.8000	73.0000
2012.01	78.2991	76.7451	79.8531	55.8000	58.2000
2012.02	66.1308	64.7239	67.5377	29.2000	33.1000
2012.03	73.7753	72.4565	75.0940	53.1000	64.1000
2012.04	78.3510	76.0409	80.6611	51.4000	55.2000
2012.05	83.8057	82.3436	85.2679	61.8000	69.0000
2012.06	68.3582	67.1488	69.5676	59.7000	64.5000
2012.07	75.6738	74.4001	76.9476	64.2000	51.3000
2012.08	74.4590	73.2146	75.7033	57.7000	63.1000
2012.09	84.7912	83.3408	86.2417	57.7000	61.5000
2012.10	81.5436	80.0029	83.0842	48.3000	53.3000
2012.11	83.8646	82.1836	85.5456	56.7000	61.4000
2012.12	75.7323	74.1392	77.3254	37.4000	40.8000
2013.01	88.1836	86.5359	89.8314	63.8000	62.9000
2013.02	76.2369	74.7503	77.7235	37.8000	38.0000
2013.03	80.9833	79.4589	82.5076	50.6000	57.9000
2013.04	91.2002	89.6655	92.7349	70.6000	72.4000
2013.05	91.4373	89.8621	93.0125	77.4000	78.7000
2013.06	75.2104	73.8749	76.5459	51.0000	52.5000
2013.07	81.2398	79.9760	82.5036	57.0000	57.0000
2013.08	82.0495	80.7687	83.3302	60.0000	66.0000
2013.09	92.5473	90.9477	94.1470	34.6000	36.9000
2013.10	87.4447	85.8942	88.9952	74.5000	85.6000
2013.11	90.0570	88.1851	91.9289	73.9000	77.6000
2013.12	83.3973	81.7211	85.0734	77.8000	90.3000
2014.01	104.7289	102.5265	106.9314	77.4000	82.0000
2014.02	90.5046	88.7851	92.2240	93.9000	102.8000
2014.03	99.6098	97.9287	101.2908	80.9000	92.2000
2014.04	110.8564	108.9752	112.7376	76.9000	84.7000
2014.05	110.7137	108.9429	112.4845	72.3000	75.2000
2014.06	91.0392	89.5636	92.5147	67.2000	71.0000
2014.07	99.6661	98.0408	101.2914	72.5000	72.5000
2014.08	100.1774	98.6732	101.6816	71.2000	74.7000
2014.09	114.1735	112.3314	116.0155	83.2000	87.6000
2014.10	107.6668	105.8658	109.4677	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	111.4176	109.3160	113.5191	65.8000	71.1000
2014.12	100.8565	98.6682	103.0449	75.8000	78.0000
2015.01	63.9101	62.6783	65.1419	65.9000	67.0000
2015.02	55.2072	53.9184	56.4960	42.4000	44.8000
2015.03	59.5511	58.4580	60.6442	38.0000	38.4000
2015.04	67.0902	65.9001	68.2802	49.0000	54.4000
2015.05	66.6091	65.5393	67.6789	56.3000	58.8000
2015.06	55.2559	54.3140	56.1979	50.2000	68.3000
2015.07	59.3031	58.2899	60.3162	47.9000	65.8000
2015.08	61.0534	60.0527	62.0542	39.5000	57.2000
2015.09	69.3058	68.1774	70.4342	49.2000	72.1000
2015.10	65.3241	64.2106	66.4375	39.3000	48.3000
2015.11	68.1038	67.1819	69.0257	39.6000	55.9000
2015.12	61.1754	59.9432	62.4076	36.4000	44.8000
2016.01	35.6519	35.0186	36.2852	33.7000	43.3000
2016.02	30.1633	29.5652	30.7614	38.3000	46.8000
2016.03	32.4183	31.8219	33.0146	30.5000	38.9000
2016.04	35.9003	35.2624	36.5383	26.6000	30.9000
2016.05	36.6282	36.0022	37.2542	33.7000	48.4000
2016.06	30.0087	29.5326	30.4849	13.1000	19.5000
2016.07	32.8929	32.3849	33.4010	21.2000	27.5000
2016.08	33.5176	32.9657	34.0696	33.0000	47.9000
2016.09	37.8122	37.1848	38.4396	27.7000	37.1000
2016.10	35.8743	35.2526	36.4959	22.7000	31.7000
2016.11	36.9875	36.3505	37.6245	14.0000	22.2000
2016.12	33.5318	32.8592	34.2044	11.1000	20.0000
2017.01	20.9517	20.5529	21.3505	18.4000	26.2000
2017.02	17.5624	17.2182	17.9066	14.4000	20.6000
2017.03	19.2990	18.9575	19.6404	11.3000	15.5000
2017.04	21.7835	21.4286	22.1384	21.6000	33.2000
2017.05	21.6594	21.3064	22.0124	12.5000	18.1000
2017.06	17.8449	17.6146	18.0751	15.5000	19.3000

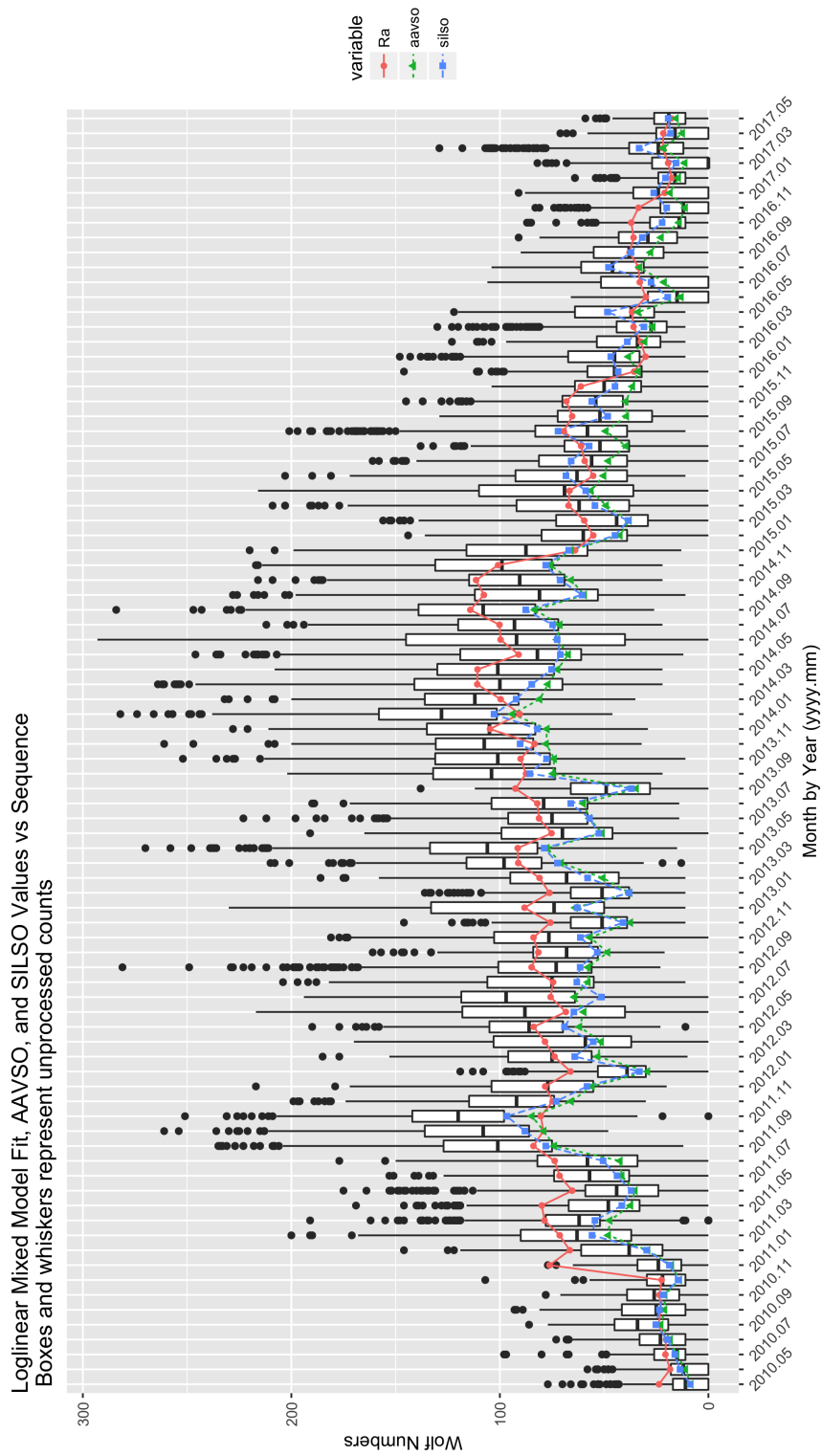


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201706 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2323	0.0324	99.7725	0.0000
seeF	-0.1886	0.0071	-26.3757	0.0000
seeG	-0.1014	0.0062	-16.3149	0.0000
seeP	-0.2928	0.0105	-27.9321	0.0000
silso1	0.1058	0.0478	2.2118	0.0270
year2011	1.2053	0.0154	78.3659	0.0000
year2012	1.2234	0.0153	79.8128	0.0000
year2013	1.3184	0.0153	86.2540	0.0000
year2014	1.5070	0.0152	99.2556	0.0000
year2015	1.0089	0.0156	64.7658	0.0000
year2016	0.4038	0.0166	24.3746	0.0000
year2017	-0.1264	0.0207	-6.1011	0.0000
mon2	-0.1555	0.0118	-13.1558	0.0000
mon3	-0.0790	0.0109	-7.2487	0.0000
mon4	0.0311	0.0109	2.8562	0.0043
mon5	0.0333	0.0104	3.2129	0.0013
mon6	-0.1714	0.0109	-15.7266	0.0000
mon7	-0.0860	0.0106	-8.1068	0.0000
mon8	-0.0702	0.0104	-6.7206	0.0000
mon9	0.0607	0.0101	6.0288	0.0000
mon10	0.0081	0.0106	0.7586	0.4481
mon11	0.0463	0.0109	4.2478	0.0000
mon12	-0.0431	0.0116	-3.7200	0.0002

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: 201706 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2553	Min. :1721096	Min. :2010	Min. : 1.00	Min. : 1.00
CHAG : 2340	1st Qu.:2456060	1st Qu.:2012	1st Qu.: 4.00	1st Qu.: 8.00
BRAB : 2335	Median :2456646	Median :2013	Median : 7.00	Median :16.00
BROB : 2065	Mean :2456364	Mean :2013	Mean : 6.57	Mean :15.72
KNJS : 1950	3rd Qu.:2457281	3rd Qu.:2015	3rd Qu.: 9.00	3rd Qu.:23.00
HOWR : 1938	Max. :2457935	Max. :2017	Max. :12.00	Max. :31.00
(Other):46664				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:10935	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :24894	Min. :0.0000
F:18423	1st Qu.: 2.000	1st Qu.: 8.00	1st Qu.: 32.00	3000F : 9764	1st Qu.:0.0000
G:25556	Median : 4.000	Median : 19.00	Median : 60.00	2500E : 7766	Median :0.0000
P: 4931	Mean : 4.105	Mean : 25.45	Mean : 66.51	3500G : 4618	Mean :0.3288
	3rd Qu.: 6.000	3rd Qu.: 37.00	3rd Qu.: 95.00	1000B : 4228	3rd Qu.:1.0000
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3059	Max. :1.0000
				(Other): 5516	

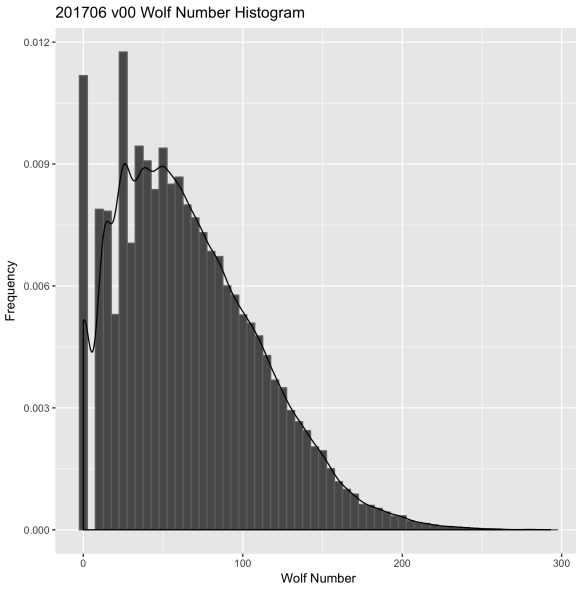


Figure 3: Box plots of raw Wolf number (w) by observer rank.

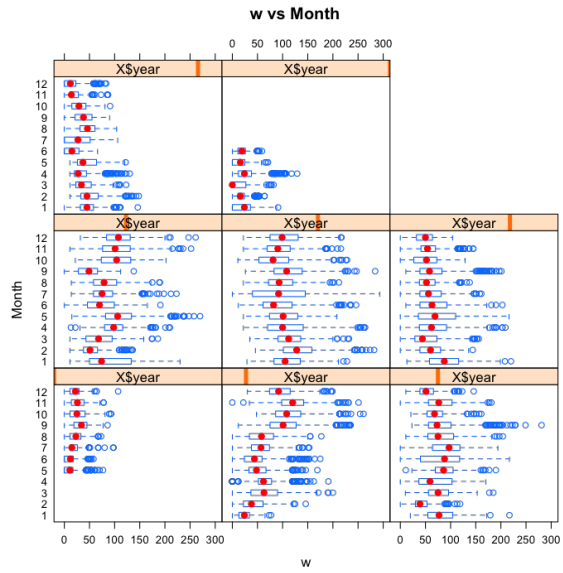


Figure 4: Box plots of raw Wolf number (w) by month and year.

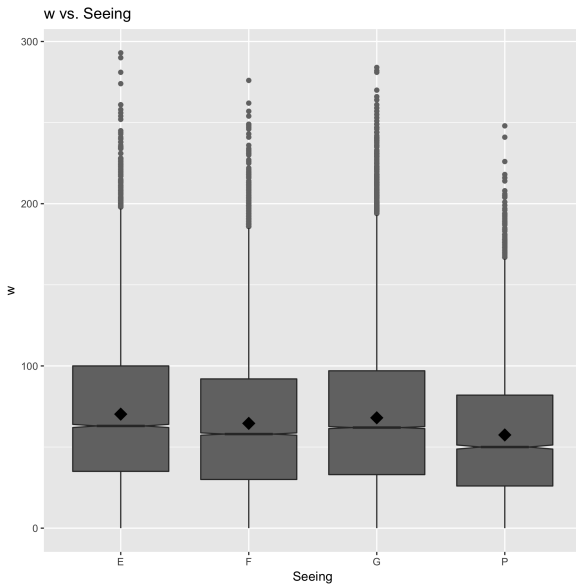


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

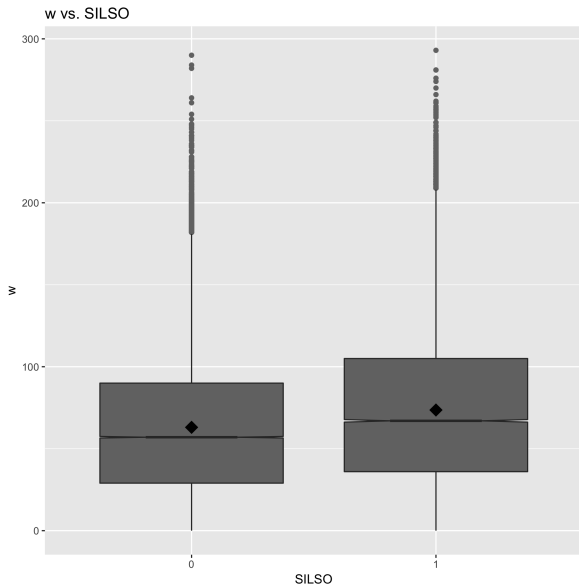


Figure 6: Box plots of raw Wolf number (w) by organization.

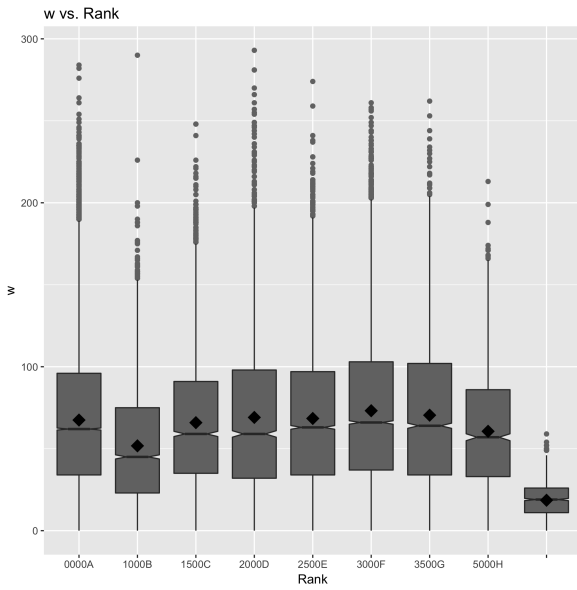


Figure 7: Box plots of raw Wolf number (w) by observer rank (deprecated 20170713 for further development).

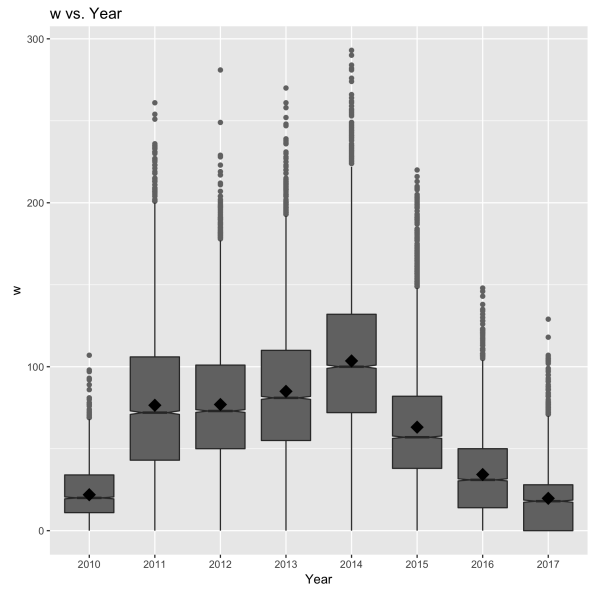


Figure 8: Box plots of raw Wolf number (w) by year.