

Monthly Report (00)

2016.12 Data Set

Friday 13th January, 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.

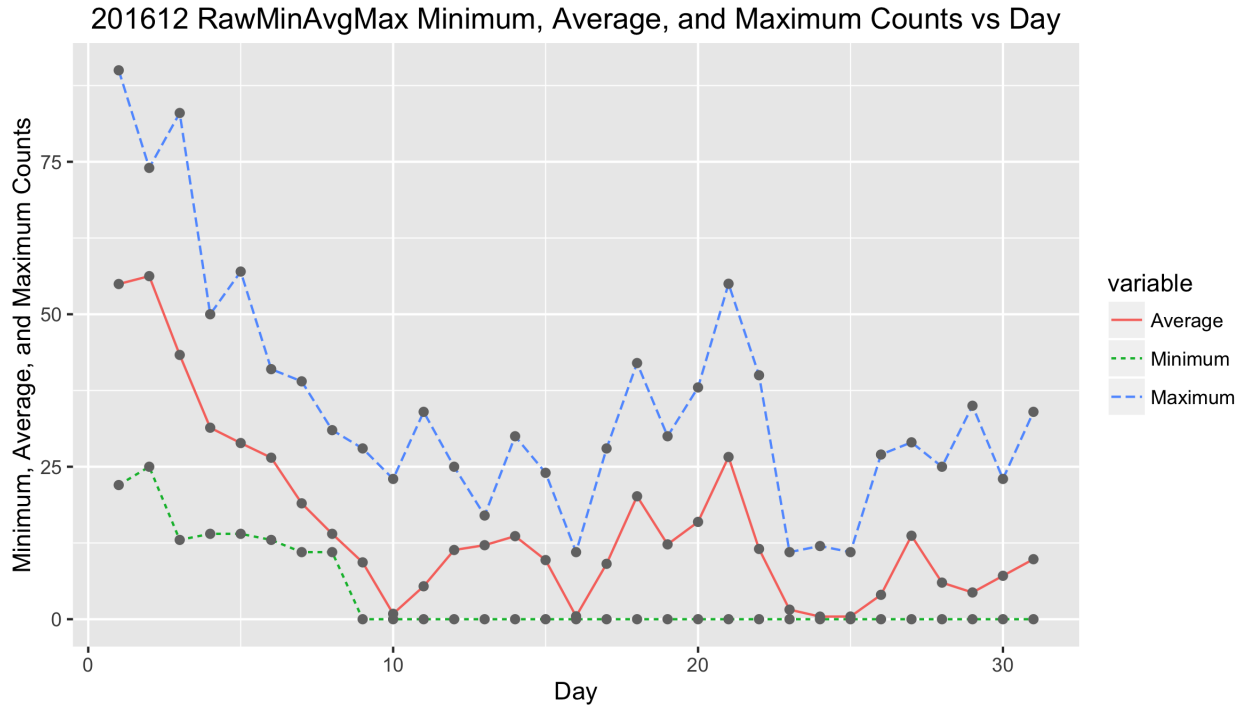


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201612 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	27.0000	22.0000	54.9583	90.0000
2.0000	27.0000	25.0000	56.2500	74.0000
3.0000	29.0000	13.0000	43.3333	83.0000
4.0000	27.0000	14.0000	31.4000	50.0000
5.0000	28.0000	14.0000	28.8846	57.0000
6.0000	32.0000	13.0000	26.4815	41.0000
7.0000	27.0000	11.0000	19.0000	39.0000
8.0000	29.0000	11.0000	14.0000	31.0000
9.0000	29.0000	0.0000	9.3200	28.0000
10.0000	32.0000	0.0000	0.8846	23.0000
11.0000	29.0000	0.0000	5.4000	34.0000
12.0000	25.0000	0.0000	11.3478	25.0000
13.0000	22.0000	0.0000	12.1364	17.0000
14.0000	23.0000	0.0000	13.6190	30.0000
15.0000	25.0000	0.0000	9.6957	24.0000
16.0000	23.0000	0.0000	0.4783	11.0000
17.0000	26.0000	0.0000	9.0833	28.0000
18.0000	33.0000	0.0000	20.1538	42.0000
19.0000	29.0000	0.0000	12.2692	30.0000
20.0000	27.0000	0.0000	15.9565	38.0000
21.0000	22.0000	0.0000	26.6000	55.0000
22.0000	27.0000	0.0000	11.5200	40.0000
23.0000	31.0000	0.0000	1.5714	11.0000
24.0000	33.0000	0.0000	0.4286	12.0000
25.0000	31.0000	0.0000	0.4400	11.0000
26.0000	28.0000	0.0000	4.0000	27.0000
27.0000	36.0000	0.0000	13.6875	29.0000
28.0000	29.0000	0.0000	6.0000	25.0000
29.0000	28.0000	0.0000	4.3846	35.0000
30.0000	30.0000	0.0000	7.1200	23.0000
31.0000	32.0000	0.0000	9.8400	34.0000

3 Error Tables

Data are for the month of December 2016. 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	23.8963	23.3488	24.4438	8.4000	8.7000
2010.06	18.3068	17.8179	18.7956	11.0000	13.6000
2010.07	20.5947	20.1389	21.0505	15.2000	16.1000
2010.08	20.3033	19.8085	20.7981	18.3000	19.6000
2010.09	23.8866	23.3731	24.4001	22.8000	25.2000
2010.10	22.6984	22.2081	23.1887	21.0000	23.5000
2010.11	23.4225	22.8927	23.9522	20.9000	21.6000
2010.12	22.4200	21.7729	23.0670	13.9000	14.5000
2011.01	75.8030	74.0964	77.5095	17.7000	18.7000
2011.02	66.2946	64.7985	67.7908	29.1000	29.6000
2011.03	71.9894	70.5019	73.4769	48.0000	55.8000
2011.04	77.5538	75.8938	79.2138	47.3000	54.4000
2011.05	80.6775	79.0659	82.2892	37.3000	41.5000
2011.06	65.1253	63.7525	66.4981	35.2000	37.0000
2011.07	71.3887	69.8155	72.9619	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	73.7287	72.2819	75.1755	42.4000	50.5000
2011.09	83.9413	82.8358	85.0468	73.8000	78.0000
2011.10	79.1706	77.8095	80.5318	78.9000	88.0000
2011.11	80.4108	78.6854	82.1362	84.6000	96.7000
2011.12	74.7547	73.1096	76.3998	65.8000	73.0000
2012.01	77.8511	76.3043	79.3979	55.8000	58.2000
2012.02	65.8708	64.4688	67.2727	29.2000	33.1000
2012.03	74.4215	73.0896	75.7534	53.1000	64.1000
2012.04	77.2424	74.9598	79.5250	51.4000	55.2000
2012.05	84.5921	83.1138	86.0704	61.8000	69.0000
2012.06	68.1069	66.8979	69.3160	59.7000	64.5000
2012.07	75.6131	74.3377	76.8885	64.2000	51.3000
2012.08	74.3940	73.1481	75.6400	57.7000	63.1000
2012.09	84.6975	83.2471	86.1479	57.7000	61.5000
2012.10	81.4577	79.9181	82.9972	48.3000	53.3000
2012.11	83.7793	82.0993	85.4594	56.7000	61.4000
2012.12	75.6515	74.0581	77.2449	37.4000	40.8000
2013.01	87.8066	86.1628	89.4504	63.8000	62.9000
2013.02	76.0332	74.5502	77.5161	37.8000	38.0000
2013.03	81.7633	80.2224	83.3042	50.6000	57.9000
2013.04	90.0056	88.4901	91.5212	70.6000	72.4000
2013.05	92.4270	90.8325	94.0214	77.4000	78.7000
2013.06	75.0367	73.7003	76.3731	51.0000	52.5000
2013.07	81.2479	79.9823	82.5136	57.0000	57.0000
2013.08	82.0742	80.7917	83.3568	60.0000	66.0000
2013.09	92.5647	90.9651	94.1642	34.6000	36.9000
2013.10	87.4884	85.9352	89.0417	74.5000	85.6000
2013.11	90.0770	88.2051	91.9488	73.9000	77.6000
2013.12	83.4326	81.7570	85.1083	77.8000	90.3000
2014.01	104.2657	102.0737	106.4576	77.4000	82.0000
2014.02	90.2730	88.5591	91.9869	93.9000	102.8000
2014.03	100.5319	98.8342	102.2297	80.9000	92.2000
2014.04	109.3553	107.4966	111.2140	76.9000	84.7000
2014.05	111.8641	110.0729	113.6553	72.3000	75.2000
2014.06	90.8143	89.3416	92.2870	67.2000	71.0000
2014.07	99.7117	98.0857	101.3377	72.5000	72.5000
2014.08	100.1975	98.6911	101.7038	71.2000	74.7000
2014.09	114.1861	112.3447	116.0275	83.2000	87.6000
2014.10	107.6856	105.8825	109.4887	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	111.4251	109.3236	113.5265	65.8000	71.1000
2014.12	100.8516	98.6671	103.0361	75.8000	78.0000
2015.01	63.6087	62.3819	64.8355	65.9000	67.0000
2015.02	55.0703	53.7851	56.3555	42.4000	44.8000
2015.03	60.1345	59.0303	61.2386	38.0000	38.4000
2015.04	66.2138	65.0383	67.3892	49.0000	54.4000
2015.05	67.3327	66.2514	68.4139	56.3000	58.8000
2015.06	55.1272	54.1873	56.0672	50.2000	68.3000
2015.07	59.3118	58.2975	60.3261	47.9000	65.8000
2015.08	61.0760	60.0743	62.0778	39.5000	57.2000
2015.09	69.3038	68.1744	70.4332	49.2000	72.1000
2015.10	65.3289	64.2131	66.4447	39.3000	48.3000
2015.11	68.1083	67.1854	69.0311	39.6000	55.9000
2015.12	61.1520	59.9153	62.3886	36.4000	44.8000
2016.01	35.4949	34.8649	36.1248	33.7000	43.3000
2016.02	30.0764	29.4796	30.6732	38.3000	46.8000
2016.03	32.7312	32.1306	33.3319	30.5000	38.9000
2016.04	35.4229	34.7935	36.0522	26.6000	30.9000
2016.05	37.0135	36.3797	37.6473	33.7000	48.4000
2016.06	29.9360	29.4607	30.4113	13.1000	19.5000
2016.07	32.9038	32.3949	33.4126	21.2000	27.5000
2016.08	33.5258	32.9726	34.0789	33.0000	47.9000
2016.09	37.8136	37.1854	38.4419	27.7000	37.1000
2016.10	35.8909	35.2692	36.5125	22.7000	31.7000
2016.11	37.0075	36.3704	37.6445	14.0000	22.2000
2016.12	33.5497	32.8765	34.2229	11.1000	20.0000

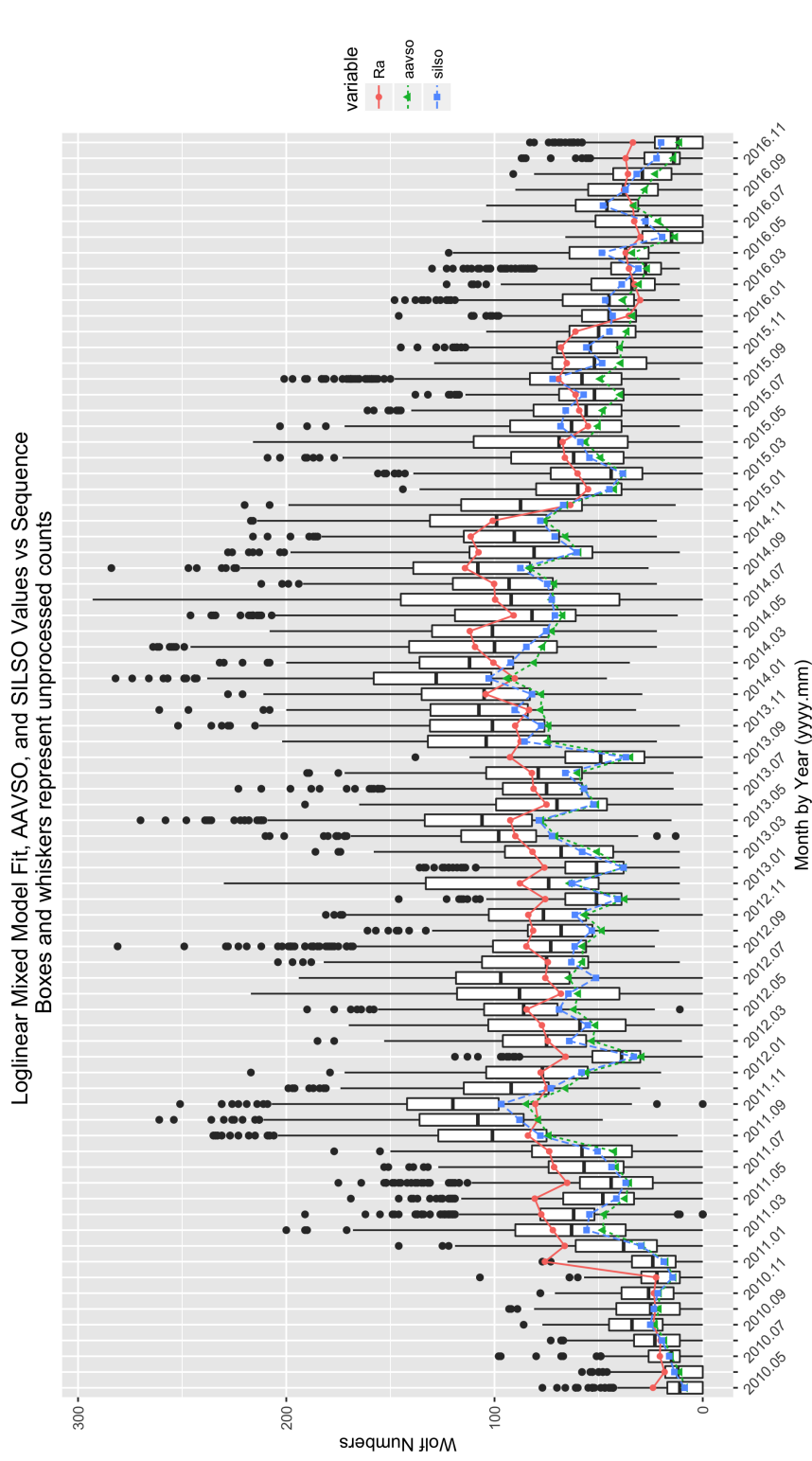


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201612 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2153	0.0445	72.2407	0.0000
seeF	-0.1889	0.0072	-26.1252	0.0000
seeG	-0.1024	0.0063	-16.2983	0.0000
seeP	-0.2954	0.0106	-27.8417	0.0000
r1000B	-0.0580	0.0831	-0.6977	0.4854
r1500C	0.0366	0.1270	0.2883	0.7731
r2000D	0.0768	0.1549	0.4961	0.6198
r2500E	-0.0011	0.1052	-0.0104	0.9917
r3000F	0.0726	0.1024	0.7093	0.4781
r3500G	0.1221	0.1533	0.7966	0.4257
r5000H	-0.1052	0.2121	-0.4959	0.6199
silsoy	0.1219	0.0738	1.6513	0.0987
year2011	1.2060	0.0154	78.3489	0.0000
year2012	1.2226	0.0153	79.6866	0.0000
year2013	1.3188	0.0153	86.2070	0.0000
year2014	1.5073	0.0152	99.1854	0.0000
year2015	1.0093	0.0156	64.7349	0.0000
year2016	0.4043	0.0166	24.3815	0.0000
mon2	-0.1537	0.0121	-12.6996	0.0000
mon3	-0.0649	0.0111	-5.8325	0.0000
mon4	0.0222	0.0112	1.9882	0.0468
mon5	0.0484	0.0106	4.5709	0.0000
mon6	-0.1695	0.0112	-15.1345	0.0000
mon7	-0.0812	0.0107	-7.5589	0.0000
mon8	-0.0653	0.0106	-6.1736	0.0000
mon9	0.0653	0.0102	6.3985	0.0000
mon10	0.0128	0.0108	1.1831	0.2368
mon11	0.0510	0.0110	4.6206	0.0000
mon12	-0.0384	0.0117	-3.2787	0.0010

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: 201612 Summary of Sunspot Numbers

obs	jd	year	mon	day
ARAG : 2375	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2183	1st Qu.:2455999	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2160	Median :2456546	Median :2013	Median : 7.000	Median :16.00
BROB : 1928	Mean :2456247	Mean :2013	Mean : 6.775	Mean :15.72
HOWR : 1796	3rd Qu.:2457153	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
KNJS : 1778	Max. :2457754	Max. :2016	Max. :12.000	Max. :31.00
(Other):43284				

Table 5: Summary of Sunspot Numbers

see	g	s	w	r	silso
E:10166	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :23605	n:37320
F:17021	1st Qu.: 2.000	1st Qu.: 9.00	1st Qu.: 37.00	3000F : 9160	y:18184
G:23740	Median : 4.000	Median : 21.00	Median : 64.00	2500E : 7276	
P: 4577	Mean : 4.322	Mean : 26.94	Mean : 70.17	3500G : 4343	
	3rd Qu.: 6.000	3rd Qu.: 38.00	3rd Qu.: 98.00	1000B : 4028	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3021	
				(Other): 4071	

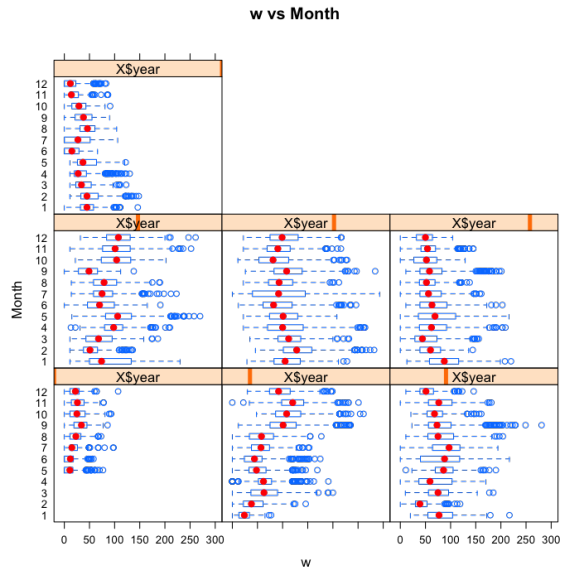
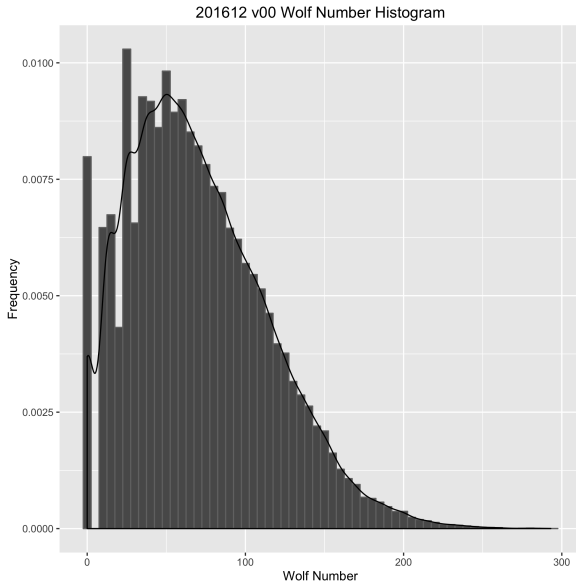


Figure 3: Box plots of raw Wolf number (w) by observer rank.

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

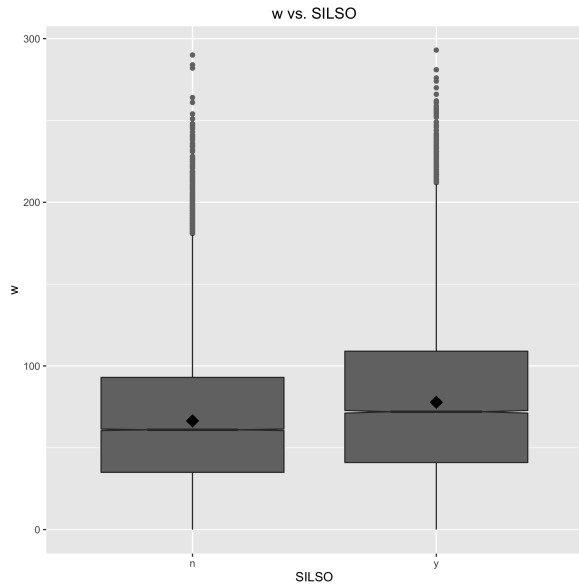
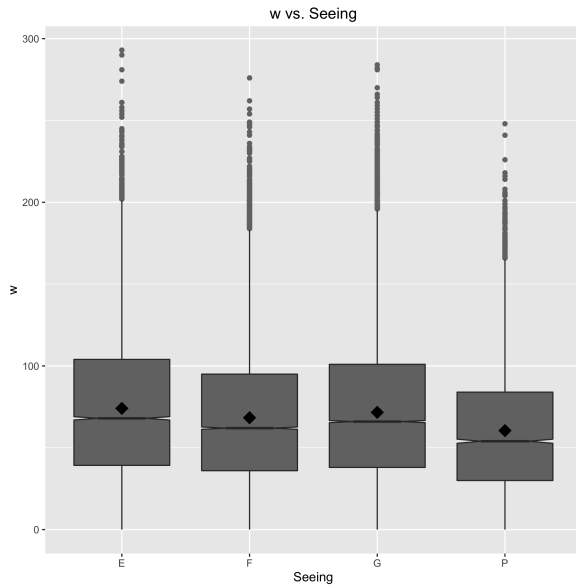


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

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

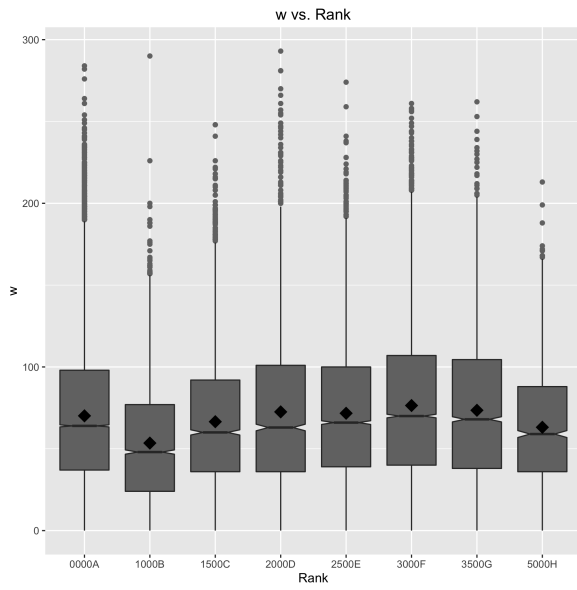


Figure 7: Box plots of raw Wolf number (w) by observer rank.

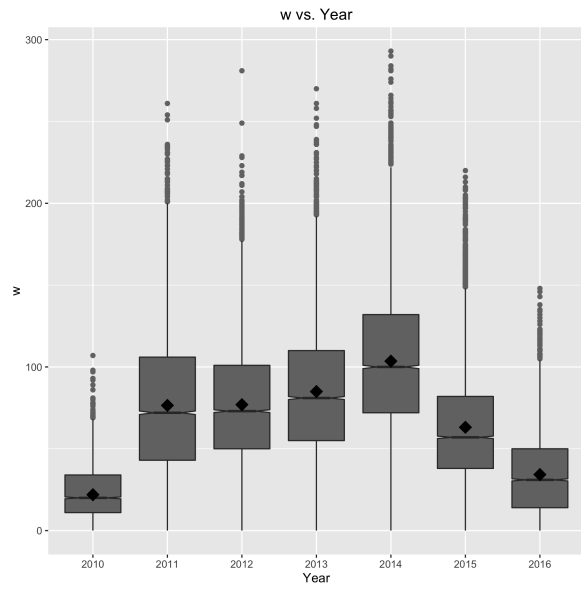


Figure 8: Box plots of raw Wolf number (w) by year.