Monthly Report (00)

2016.09 Data Set

Monday 17th October, 2016

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.

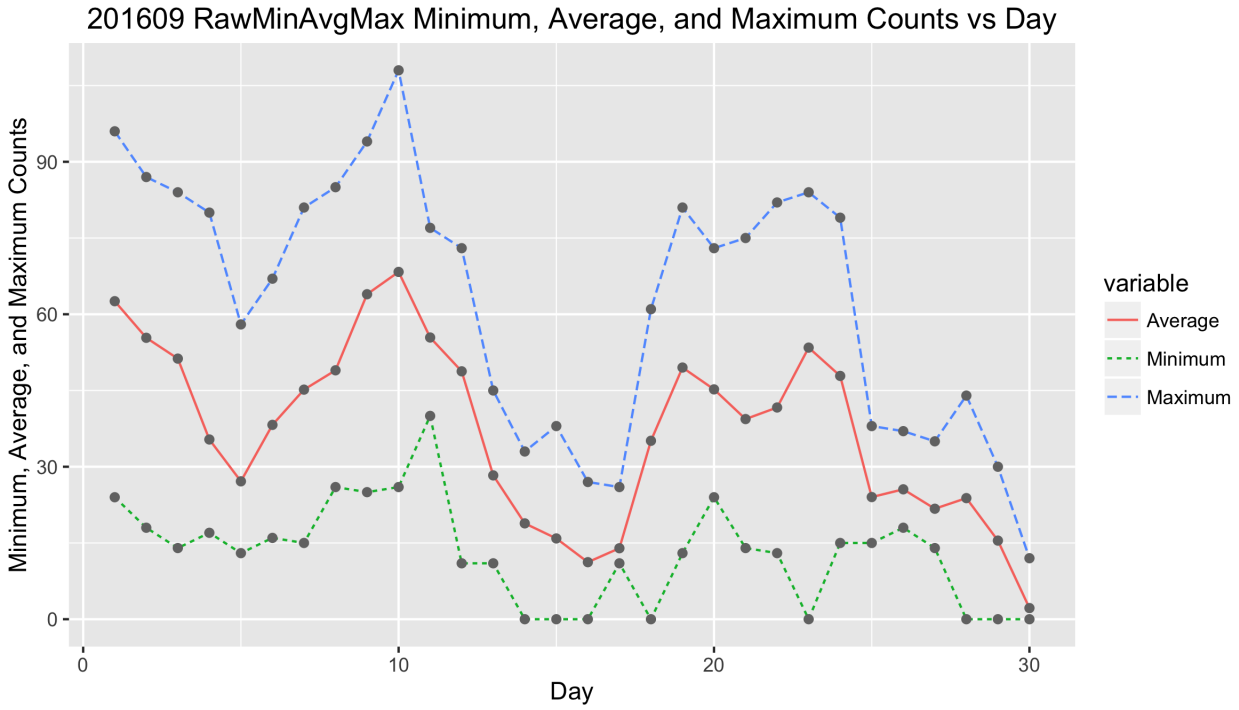


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201609 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	37.0000	24.0000	62.5714	96.0000
2.0000	41.0000	18.0000	55.3684	87.0000
3.0000	35.0000	14.0000	51.2727	84.0000
4.0000	39.0000	17.0000	35.3611	80.0000
5.0000	33.0000	13.0000	27.1290	58.0000
6.0000	37.0000	16.0000	38.2353	67.0000
7.0000	36.0000	15.0000	45.1714	81.0000
8.0000	39.0000	26.0000	48.9714	85.0000
9.0000	37.0000	25.0000	63.9429	94.0000
10.0000	42.0000	26.0000	68.3333	108.0000
11.0000	40.0000	40.0000	55.4242	77.0000
12.0000	39.0000	11.0000	48.7714	73.0000
13.0000	30.0000	11.0000	28.2963	45.0000
14.0000	31.0000	0.0000	18.8571	33.0000
15.0000	29.0000	0.0000	15.8929	38.0000
16.0000	29.0000	0.0000	11.2222	27.0000
17.0000	33.0000	11.0000	13.9655	26.0000
18.0000	30.0000	0.0000	35.1071	61.0000
19.0000	31.0000	13.0000	49.5172	81.0000
20.0000	28.0000	24.0000	45.2222	73.0000
21.0000	30.0000	14.0000	39.3793	75.0000
22.0000	31.0000	13.0000	41.6552	82.0000
23.0000	34.0000	0.0000	53.4375	84.0000
24.0000	31.0000	15.0000	47.8667	79.0000
25.0000	33.0000	15.0000	24.0333	38.0000
26.0000	32.0000	18.0000	25.5556	37.0000
27.0000	36.0000	14.0000	21.7353	35.0000
28.0000	31.0000	0.0000	23.8276	44.0000
29.0000	32.0000	0.0000	15.4667	30.0000
30.0000	27.0000	0.0000	2.1923	12.0000

3 Error Tables

Data are for the month of September 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 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
2010.0500	23.5887	23.0503	24.1270	8.4000	8.7000
2010.0600	18.0540	17.5722	18.5358	11.0000	13.6000
2010.0700	20.3234	19.8748	20.7720	15.2000	16.1000
2010.0800	20.0474	19.5593	20.5355	18.3000	19.6000
2010.0900	23.6213	23.1150	24.1276	22.8000	25.2000
2010.1000	22.8087	22.3163	23.3012	21.0000	23.5000
2010.1100	24.1290	23.5839	24.6740	20.9000	21.6000
2010.1200	23.3492	22.6772	24.0212	13.9000	14.5000
2011.0100	75.1015	73.4146	76.7884	17.7000	18.7000
2011.0200	65.6827	64.2009	67.1645	29.1000	29.6000
2011.0300	71.3554	69.8813	72.8294	48.0000	55.8000
2011.0400	76.7302	75.0877	78.3726	47.3000	54.4000
2011.0500	79.9602	78.3632	81.5573	37.3000	41.5000
2011.0600	64.4770	63.1196	65.8343	35.2000	37.0000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	silso
2011.0700	70.7386	69.1834	72.2937	41.5000	43.8000
2011.0800	73.1057	71.6737	74.5376	42.4000	50.5000
2011.0900	83.3834	82.2878	84.4789	73.8000	78.0000
2011.1000	79.9534	78.5812	81.3256	78.9000	88.0000
2011.1100	83.1811	81.3999	84.9623	84.6000	96.7000
2011.1200	78.1346	76.4228	79.8464	65.8000	73.0000
2012.0100	77.1203	75.5898	78.6508	55.8000	58.2000
2012.0200	65.2397	63.8553	66.6240	29.2000	33.1000
2012.0300	73.8016	72.4814	75.1218	53.1000	64.1000
2012.0400	76.4028	74.1493	78.6563	51.4000	55.2000
2012.0500	83.8923	82.4251	85.3595	61.8000	69.0000
2012.0600	67.4805	66.2805	68.6805	59.7000	64.5000
2012.0700	74.9852	73.7213	76.2490	64.2000	51.3000
2012.0800	73.8355	72.6001	75.0710	57.7000	63.1000
2012.0900	84.1493	82.7104	85.5882	57.7000	61.5000
2012.1000	82.2190	80.6643	83.7737	48.3000	53.3000
2012.1100	86.6685	84.9293	88.4077	56.7000	61.4000
2012.1200	79.1266	77.4561	80.7971	37.4000	40.8000
2013.0100	87.0677	85.4308	88.7045	63.8000	62.9000
2013.0200	75.3737	73.9037	76.8437	37.8000	38.0000
2013.0300	81.1174	79.5898	82.6449	50.6000	57.9000
2013.0400	89.1219	87.6188	90.6250	70.6000	72.4000
2013.0500	91.7082	90.1227	93.2937	77.4000	78.7000
2013.0600	74.3878	73.0608	75.7147	51.0000	52.5000
2013.0700	80.5909	79.3358	81.8459	57.0000	57.0000
2013.0800	81.4829	80.2095	82.7562	60.0000	66.0000
2013.0900	92.0117	90.4200	93.6033	34.6000	36.9000
2013.1000	88.4149	86.8410	89.9889	74.5000	85.6000
2013.1100	93.2835	91.3388	95.2282	73.9000	77.6000
2013.1200	87.3890	85.6288	89.1493	77.8000	90.3000
2014.0100	103.5353	101.3564	105.7143	77.4000	82.0000
2014.0200	89.5990	87.8921	91.3059	93.9000	102.8000
2014.0300	99.7700	98.0851	101.4550	80.9000	92.2000
2014.0400	108.3427	106.4995	110.1859	76.9000	84.7000
2014.0500	111.0883	109.3112	112.8654	72.3000	75.2000
2014.0600	90.0788	88.6153	91.5423	67.2000	71.0000
2014.0700	98.9712	97.3557	100.5866	72.5000	72.5000
2014.0800	99.4750	97.9807	100.9692	71.2000	74.7000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	silso
2014.0900	113.5193	111.6921	115.3465	83.2000	87.6000
2014.1000	108.8692	107.0423	110.6961	59.5000	60.6000
2014.1100	115.4458	113.2633	117.6284	65.8000	71.1000
2014.1200	105.5698	103.2878	107.8518	75.8000	78.0000
2015.0100	62.9507	61.7420	64.1594	65.9000	67.0000
2015.0200	54.4746	53.2081	55.7411	42.4000	44.8000
2015.0300	59.5361	58.4464	60.6259	38.0000	38.4000
2015.0400	65.4363	64.2769	66.5957	49.0000	54.4000
2015.0500	66.6598	65.5900	67.7297	56.3000	58.8000
2015.0600	54.5476	53.6217	55.4735	50.2000	68.3000
2015.0700	58.6751	57.6747	59.6755	47.9000	65.8000
2015.0800	60.4387	59.4509	61.4264	39.5000	57.2000
2015.0900	68.7019	67.5834	69.8204	49.2000	72.1000
2015.1000	65.8293	64.7044	66.9541	39.3000	48.3000
2015.1100	70.3326	69.3784	71.2867	39.6000	55.9000
2015.1200	63.8206	62.5283	65.1128	36.4000	44.8000
2016.0100	39.4110	38.7119	40.1101	33.7000	43.3000
2016.0200	33.3830	32.7212	34.0449	38.3000	46.8000
2016.0300	36.3498	35.6843	37.0152	30.5000	38.9000
2016.0400	39.2866	38.5896	39.9837	26.6000	30.9000
2016.0500	41.1049	40.4008	41.8090	33.7000	48.4000
2016.0600	33.2493	32.7215	33.7771	13.1000	19.5000
2016.0700	36.5527	35.9877	37.1178	21.2000	27.5000
2016.0800	37.2637	36.6488	37.8787	33.0000	47.9000
2016.0900	42.1066	41.4066	42.8067	27.7000	37.1000

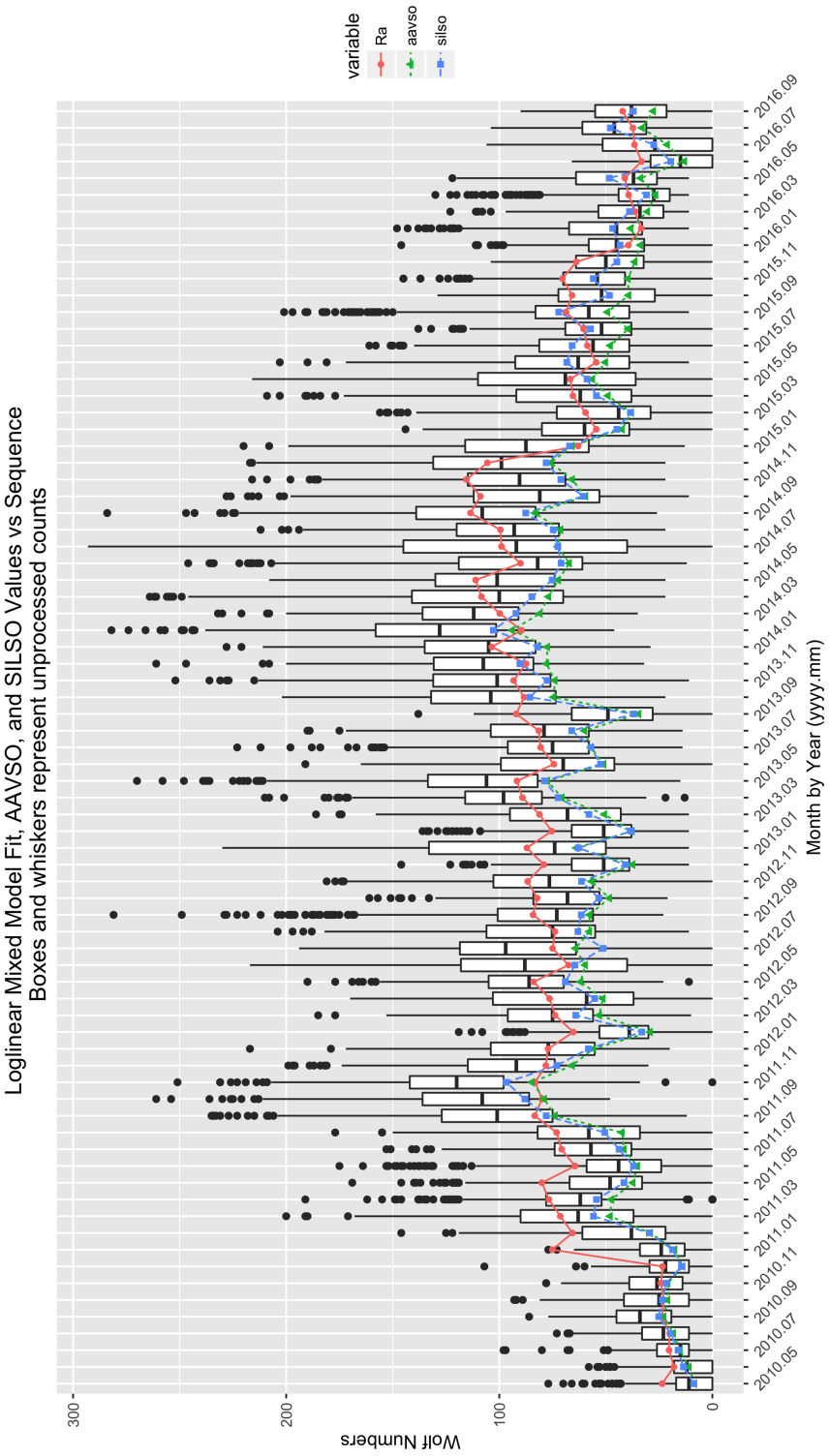


Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201609 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.2018	0.0443	72.3269	0.0000
seeF	-0.1893	0.0072	-26.2407	0.0000
seeG	-0.1024	0.0063	-16.3378	0.0000
seeP	-0.2968	0.0106	-28.0085	0.0000
r1000B	-0.0559	0.0827	-0.6757	0.4992
r1500C	0.0337	0.1264	0.2667	0.7897
r2000D	0.0791	0.1542	0.5133	0.6077
r2500E	0.0008	0.1047	0.0076	0.9939
r3000F	0.0713	0.1019	0.6998	0.4840
r3500G	0.1222	0.1526	0.8008	0.4233
r5000H	-0.1072	0.2110	-0.5080	0.6115
silsoy	0.1211	0.0735	1.6490	0.0992
year2011	1.2107	0.0153	79.1971	0.0000
year2012	1.2272	0.0152	80.5359	0.0000
year2013	1.3244	0.0152	87.1634	0.0000
year2014	1.5136	0.0151	100.2806	0.0000
year2015	1.0130	0.0155	65.4215	0.0000
year2016	0.5237	0.0169	31.0425	0.0000
mon2	-0.1536	0.0120	-12.7834	0.0000
mon3	-0.0646	0.0110	-5.8440	0.0000
mon4	0.0210	0.0111	1.8860	0.0593
mon5	0.0488	0.0105	4.6405	0.0000
mon6	-0.1701	0.0111	-15.2904	0.0000
mon7	-0.0813	0.0107	-7.6171	0.0000
mon8	-0.0648	0.0105	-6.1626	0.0000
mon9	0.0678	0.0101	6.6908	0.0000
mon10	0.0318	0.0109	2.9249	0.0034
mon11	0.0945	0.0111	8.5463	0.0000
mon12	0.0159	0.0118	1.3498	0.1771

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.

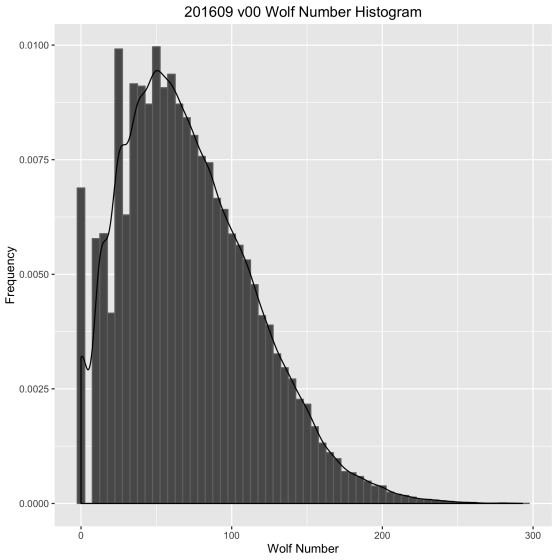
5 Supporting Information

Table 4: 201609 Summary of Sunspot Numbers

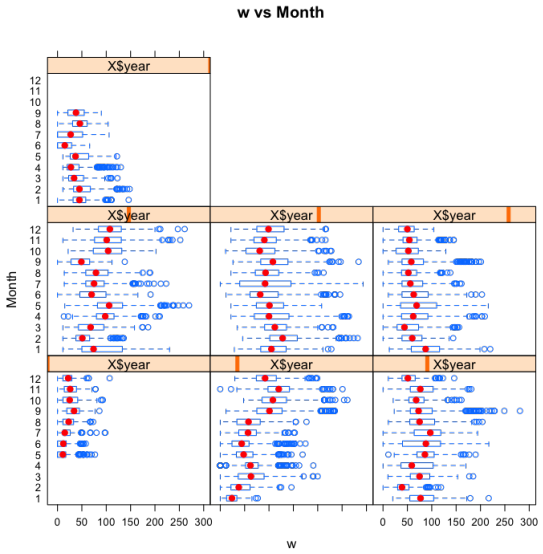
obs	jd	year	mon	day
ARAG : 2290	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2107	1st Qu.:2455979	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2070	Median :2456517	Median :2013	Median : 7.000	Median :16.00
BROB : 1856	Mean :2456200	Mean :2013	Mean : 6.639	Mean :15.72
DUBF : 1750	3rd Qu.:2457100	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 1726	Max. :2457662	Max. :2016	Max. :12.000	Max. :31.00
(Other):41958				

Table 5: Summary of Sunspot Numbers

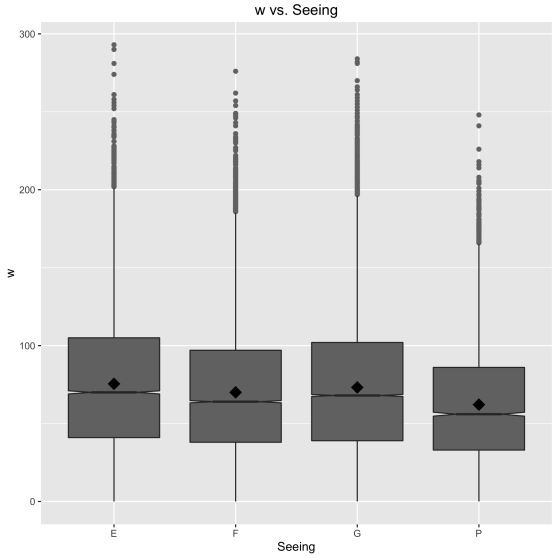
see	g	s	w	r	silso
E: 9899	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :22914	n:36203
F:16446	1st Qu.: 3.000	1st Qu.: 10.00	1st Qu.: 38.00	3000F : 8793	y:17554
G:23016	Median : 4.000	Median : 22.00	Median : 66.00	2500E : 7003	
P: 4396	Mean : 4.415	Mean : 27.58	Mean : 71.73	3500G : 4177	
	3rd Qu.: 6.000	3rd Qu.: 39.00	3rd Qu.:100.00	1000B : 3931	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3021	
				(Other): 3918	



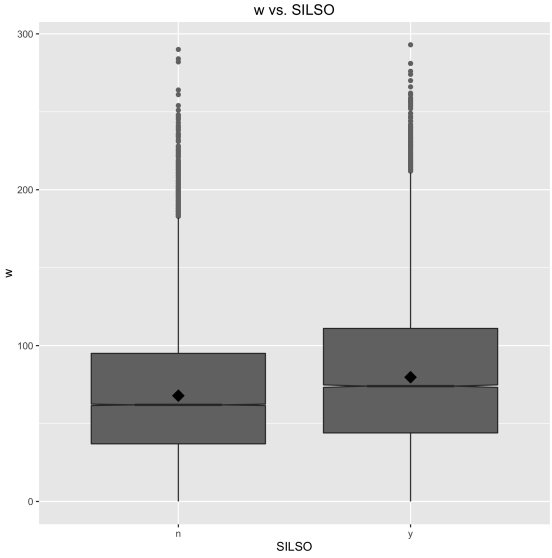
(a) Observed sunspot count histogram.



(b) Box plot of sunspot count by year and month.



(c) Box plot of sunspot count by seeing condition.



(d) Box plot of sunspot count submitted to AAVSO and SILSO.

Figure 3:

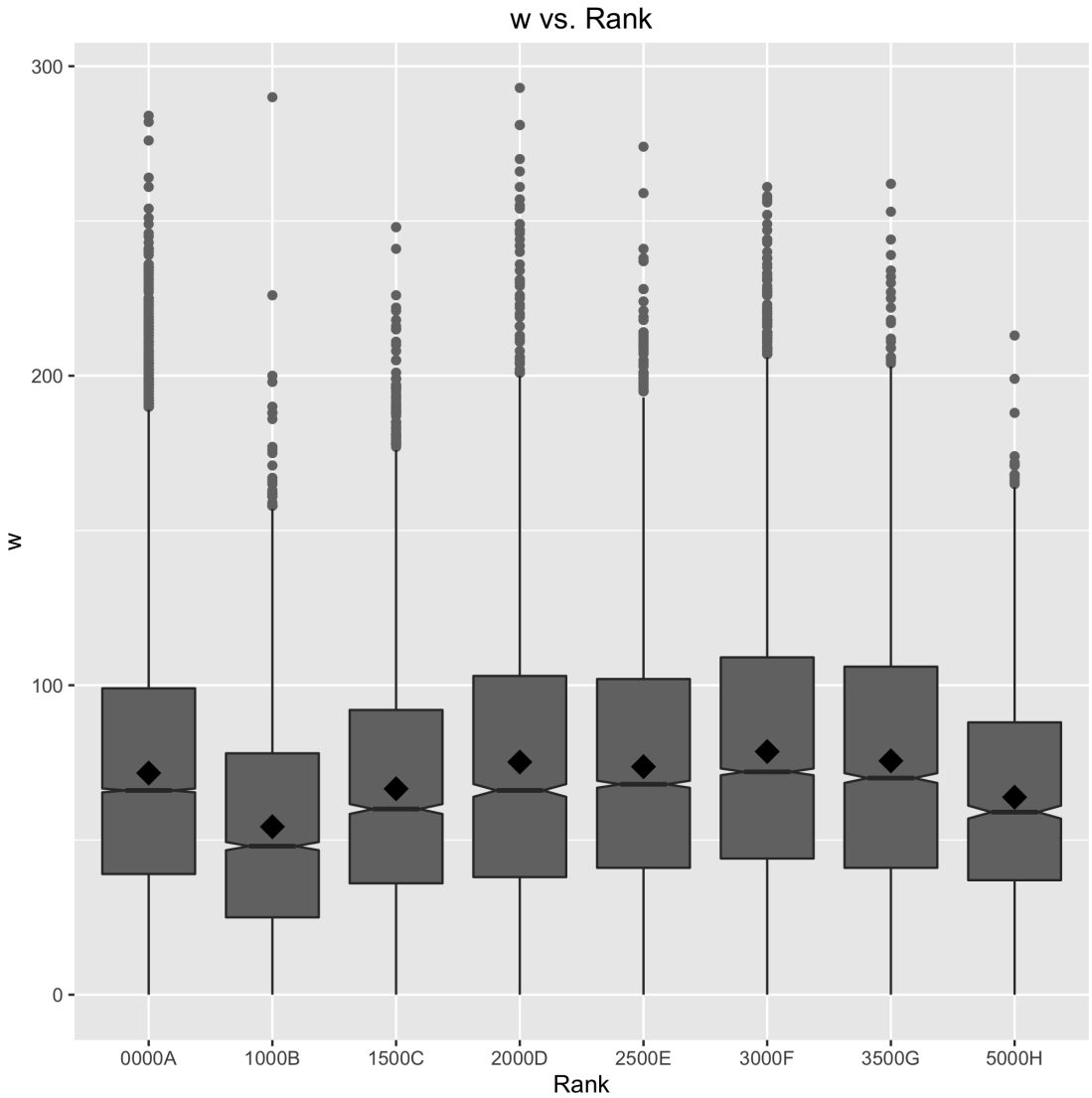


Figure 4: Box plot of sunspot count by rank.