

# Monthly Report (00)

## 2016.10 Data Set

Monday 21<sup>st</sup> November, 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.

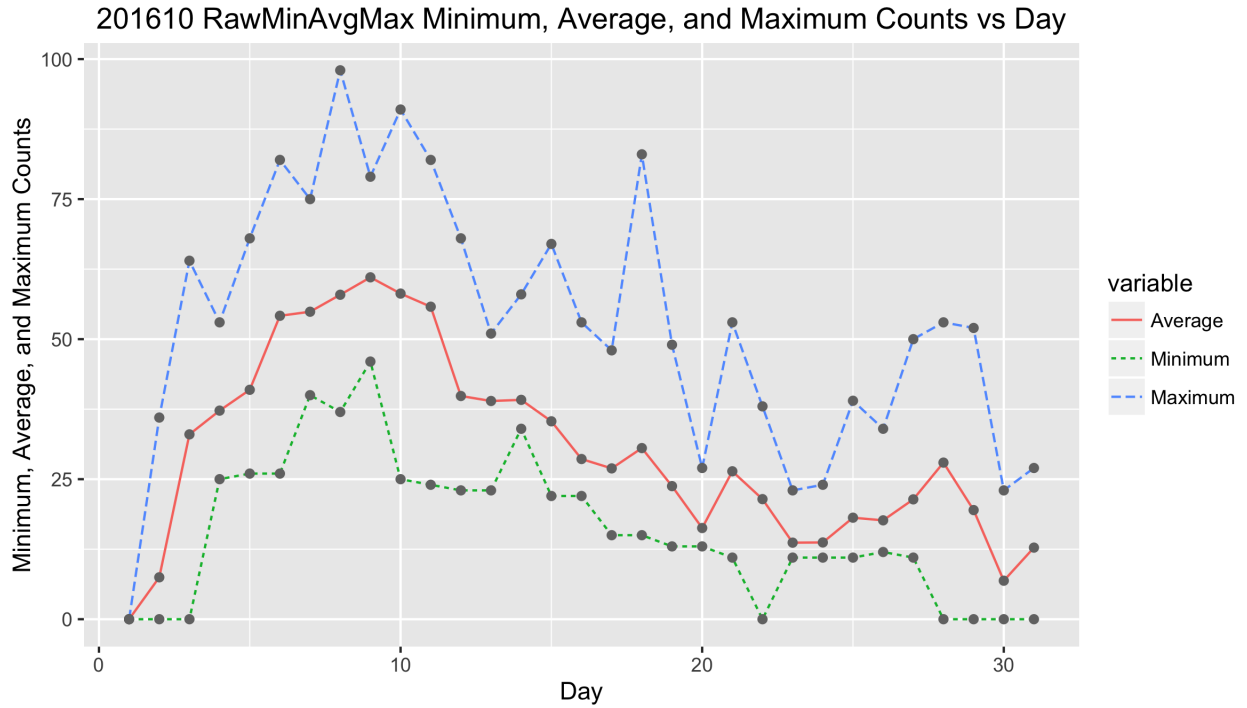


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201610 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	25.0000	0.0000	0.0000	0.0000
2.0000	32.0000	0.0000	7.4815	36.0000
3.0000	28.0000	0.0000	33.0000	64.0000
4.0000	36.0000	25.0000	37.2424	53.0000
5.0000	28.0000	26.0000	40.9600	68.0000
6.0000	31.0000	26.0000	54.1724	82.0000
7.0000	31.0000	40.0000	54.8966	75.0000
8.0000	35.0000	37.0000	57.9310	98.0000
9.0000	35.0000	46.0000	61.0385	79.0000
10.0000	35.0000	25.0000	58.1379	91.0000
11.0000	30.0000	24.0000	55.7931	82.0000
12.0000	29.0000	23.0000	39.8519	68.0000
13.0000	28.0000	23.0000	38.9600	51.0000
14.0000	34.0000	34.0000	39.1613	58.0000
15.0000	34.0000	22.0000	35.3333	67.0000
16.0000	35.0000	22.0000	28.6000	53.0000
17.0000	33.0000	15.0000	26.9259	48.0000
18.0000	30.0000	15.0000	30.5517	83.0000
19.0000	32.0000	13.0000	23.7586	49.0000
20.0000	33.0000	13.0000	16.2812	27.0000
21.0000	29.0000	11.0000	26.4074	53.0000
22.0000	30.0000	0.0000	21.4400	38.0000
23.0000	29.0000	11.0000	13.6667	23.0000
24.0000	30.0000	11.0000	13.7037	24.0000
25.0000	27.0000	11.0000	18.1250	39.0000
26.0000	25.0000	12.0000	17.6522	34.0000
27.0000	25.0000	11.0000	21.4000	50.0000
28.0000	25.0000	0.0000	27.9600	53.0000
29.0000	32.0000	0.0000	19.4828	52.0000
30.0000	37.0000	0.0000	6.8750	23.0000
31.0000	35.0000	0.0000	12.7812	27.0000

### 3 Error Tables

Data are for the month of October 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<sup>th</sup> through the 75<sup>th</sup> quartiles. The lower and upper whiskers extend 1.5 times the IQR below the 25<sup>th</sup> quartile, and 1.5 times the IQR above the 75<sup>th</sup> 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	23.6582	23.1187	24.1977	8.4000	8.7000
2010.06	18.1131	17.6302	18.5959	11.0000	13.6000
2010.07	20.3845	19.9347	20.8342	15.2000	16.1000
2010.08	20.1033	19.6144	20.5923	18.3000	19.6000
2010.09	23.6819	23.1748	24.1890	22.8000	25.2000
2010.10	22.4762	21.9914	22.9609	21.0000	23.5000
2010.11	24.1557	23.6106	24.7007	20.9000	21.6000
2010.12	23.3731	22.7007	24.0456	13.9000	14.5000
2011.01	75.2542	73.5640	76.9443	17.7000	18.7000
2011.02	65.8117	64.3288	67.2945	29.1000	29.6000
2011.03	71.4903	70.0140	72.9665	48.0000	55.8000
2011.04	76.9042	75.2601	78.5483	47.3000	54.4000
2011.05	80.1136	78.5165	81.7108	37.3000	41.5000
2011.06	64.6182	63.2602	65.9761	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
2011.07	70.8813	69.3240	72.4386	41.5000	43.8000
2011.08	73.2425	71.8082	74.6767	42.4000	50.5000
2011.09	83.5063	82.4094	84.6033	73.8000	78.0000
2011.10	78.7077	77.3583	80.0571	78.9000	88.0000
2011.11	83.1924	81.4113	84.9735	84.6000	96.7000
2011.12	78.1457	76.4333	79.8582	65.8000	73.0000
2012.01	77.2823	75.7495	78.8151	55.8000	58.2000
2012.02	65.3798	63.9923	66.7673	29.2000	33.1000
2012.03	73.9344	72.6128	75.2560	53.1000	64.1000
2012.04	76.5849	74.3290	78.8409	51.4000	55.2000
2012.05	84.0488	82.5809	85.5168	61.8000	69.0000
2012.06	67.6207	66.4201	68.8214	59.7000	64.5000
2012.07	75.1243	73.8596	76.3891	64.2000	51.3000
2012.08	73.9602	72.7237	75.1968	57.7000	63.1000
2012.09	84.2709	82.8316	85.7101	57.7000	61.5000
2012.10	80.9242	79.3963	82.4522	48.3000	53.3000
2012.11	86.6639	84.9268	88.4010	56.7000	61.4000
2012.12	79.1298	77.4629	80.7967	37.4000	40.8000
2013.01	87.2270	85.5923	88.8617	63.8000	62.9000
2013.02	75.5153	74.0444	76.9862	37.8000	38.0000
2013.03	81.2642	79.7359	82.7924	50.6000	57.9000
2013.04	89.3174	87.8144	90.8205	70.6000	72.4000
2013.05	91.8681	90.2828	93.4535	77.4000	78.7000
2013.06	74.5313	73.2041	75.8585	51.0000	52.5000
2013.07	80.7421	79.4877	81.9966	57.0000	57.0000
2013.08	81.6145	80.3413	82.8877	60.0000	66.0000
2013.09	92.1307	90.5396	93.7218	34.6000	36.9000
2013.10	87.0061	85.4603	88.5520	74.5000	85.6000
2013.11	93.2643	91.3242	95.2043	73.9000	77.6000
2013.12	87.3548	85.5979	89.1117	77.8000	90.3000
2014.01	103.7283	101.5480	105.9085	77.4000	82.0000
2014.02	89.7703	88.0642	91.4765	93.9000	102.8000
2014.03	99.9640	98.2783	101.6496	80.9000	92.2000
2014.04	108.5980	106.7541	110.4418	76.9000	84.7000
2014.05	111.2832	109.5050	113.0614	72.3000	75.2000
2014.06	90.2613	88.7980	91.7245	67.2000	71.0000
2014.07	99.1511	97.5355	100.7667	72.5000	72.5000
2014.08	99.6545	98.1602	101.1488	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
2014.09	113.6952	111.8668	115.5237	83.2000	87.6000
2014.10	107.1499	105.3555	108.9443	59.5000	60.6000
2014.11	115.4393	113.2620	117.6166	65.8000	71.1000
2014.12	105.5775	103.2963	107.8588	75.8000	78.0000
2015.01	63.0667	61.8562	64.2772	65.9000	67.0000
2015.02	54.5847	53.3156	55.8539	42.4000	44.8000
2015.03	59.6449	58.5539	60.7359	38.0000	38.4000
2015.04	65.5793	64.4184	66.7402	49.0000	54.4000
2015.05	66.7766	65.7067	67.8466	56.3000	58.8000
2015.06	54.6475	53.7209	55.5742	50.2000	68.3000
2015.07	58.7891	57.7881	59.7902	47.9000	65.8000
2015.08	60.5529	59.5645	61.5414	39.5000	57.2000
2015.09	68.8040	67.6861	69.9219	49.2000	72.1000
2015.10	64.7915	63.6862	65.8968	39.3000	48.3000
2015.11	70.3303	69.3780	71.2826	39.6000	55.9000
2015.12	63.8211	62.5325	65.1098	36.4000	44.8000
2016.01	38.5493	37.8663	39.2323	33.7000	43.3000
2016.02	32.6547	32.0079	33.3014	38.3000	46.8000
2016.03	35.5537	34.9033	36.2042	30.5000	38.9000
2016.04	38.4366	37.7556	39.1177	26.6000	30.9000
2016.05	40.2059	39.5185	40.8934	33.7000	48.4000
2016.06	32.5203	32.0054	33.0353	13.1000	19.5000
2016.07	35.7477	35.1965	36.2989	21.2000	27.5000
2016.08	36.4396	35.8398	37.0394	33.0000	47.9000
2016.09	41.1609	40.4785	41.8432	27.7000	37.1000
2016.10	39.0533	38.3778	39.7288	22.7000	31.7000



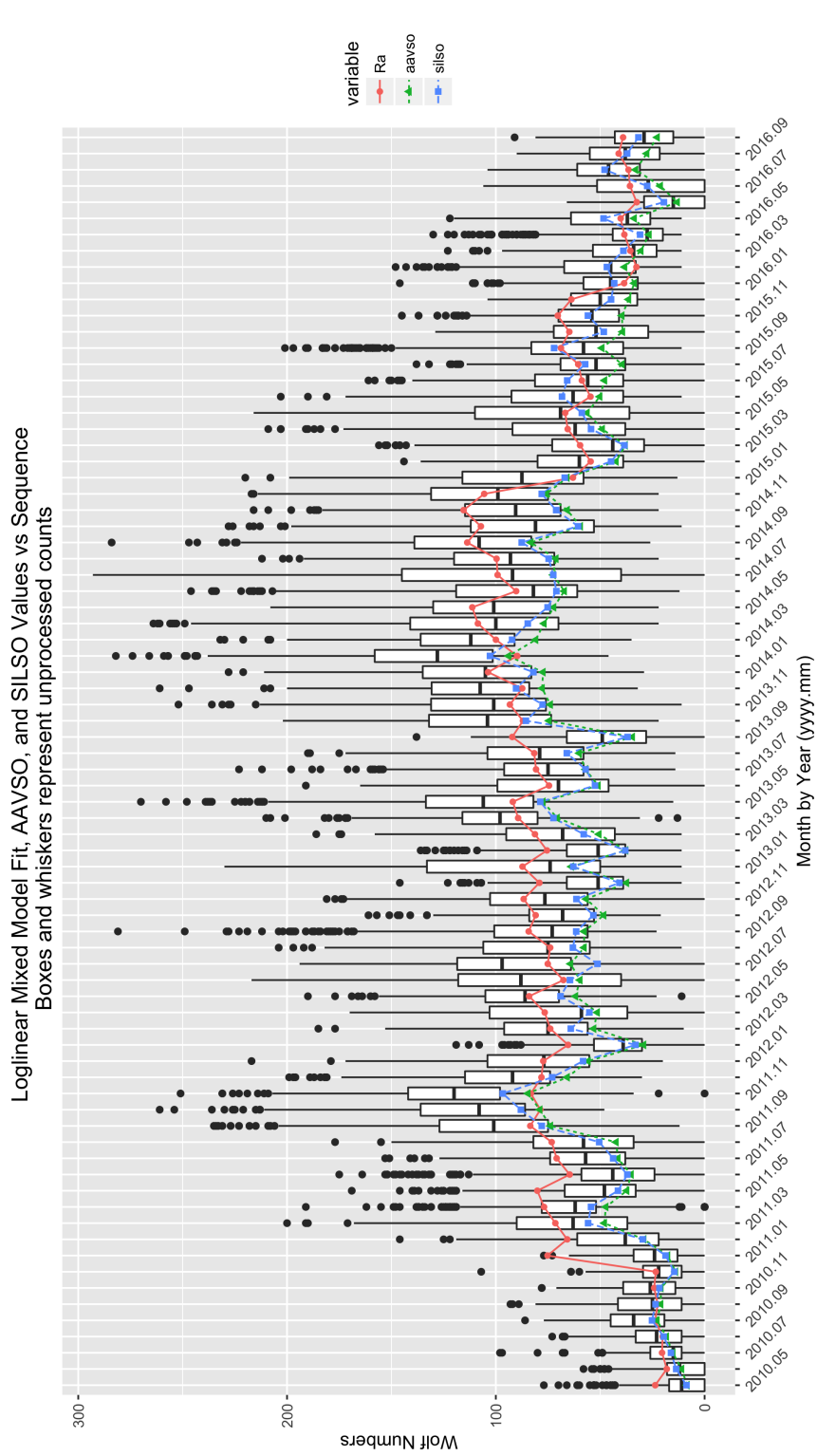


Figure 2: GLMM fitted data for  $R_a$ . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201610 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t )$
(Intercept)	3.2052	0.0443	72.4098	0.0000
seeF	-0.1891	0.0072	-26.2894	0.0000
seeG	-0.1026	0.0062	-16.4180	0.0000
seeP	-0.2966	0.0106	-28.0906	0.0000
r1000B	-0.0562	0.0827	-0.6790	0.4971
r1500C	0.0339	0.1264	0.2679	0.7888
r2000D	0.0777	0.1542	0.5038	0.6144
r2500E	0.0006	0.1047	0.0059	0.9953
r3000F	0.0713	0.1019	0.6997	0.4841
r3500G	0.1218	0.1526	0.7985	0.4246
r5000H	-0.1069	0.2110	-0.5067	0.6124
silsoy	0.1209	0.0735	1.6455	0.0999
year2011	1.2096	0.0153	79.1920	0.0000
year2012	1.2262	0.0152	80.5359	0.0000
year2013	1.3232	0.0152	87.1588	0.0000
year2014	1.5125	0.0151	100.2909	0.0000
year2015	1.0117	0.0155	65.3955	0.0000
year2016	0.4983	0.0166	29.9326	0.0000
mon2	-0.1536	0.0120	-12.7958	0.0000
mon3	-0.0646	0.0110	-5.8555	0.0000
mon4	0.0212	0.0111	1.9106	0.0561
mon5	0.0486	0.0105	4.6305	0.0000
mon6	-0.1700	0.0111	-15.2940	0.0000
mon7	-0.0813	0.0107	-7.6256	0.0000
mon8	-0.0650	0.0105	-6.1843	0.0000
mon9	0.0673	0.0101	6.6422	0.0000
mon10	0.0138	0.0107	1.2906	0.1969
mon11	0.0925	0.0111	8.3711	0.0000
mon12	0.0139	0.0118	1.1788	0.2385

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.

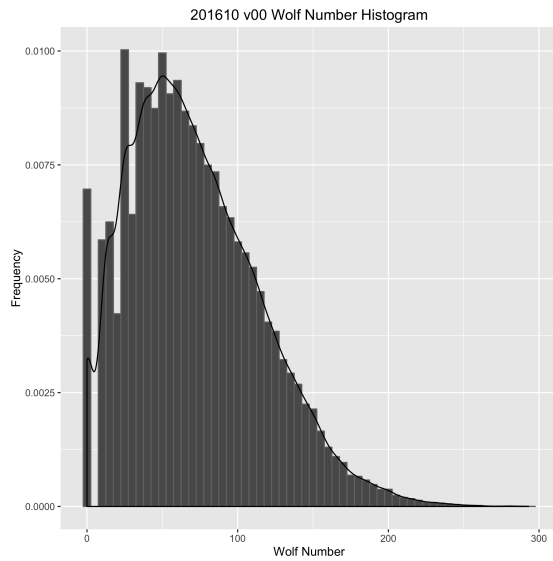
## 5 Supporting Information

Table 4: 201610 Summary of Sunspot Numbers

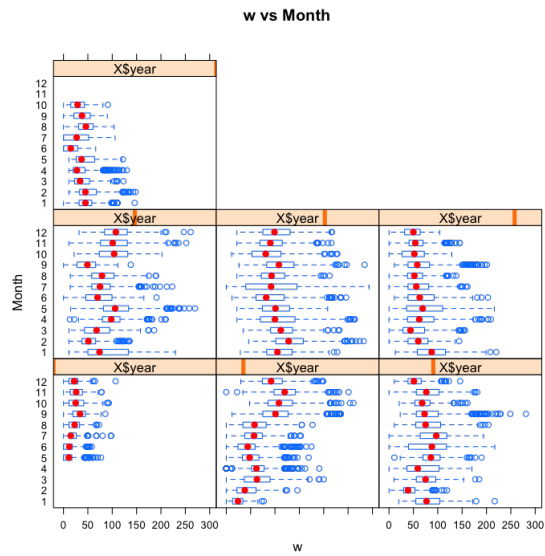
obs	jd	year	mon	day
ARAG : 2321	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2128	1st Qu.:2455989	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2100	Median :2456526	Median :2013	Median : 7.000	Median :16.00
BROB : 1878	Mean :2456217	Mean :2013	Mean : 6.678	Mean :15.72
HOWR : 1751	3rd Qu.:2457120	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
DUBF : 1750	Max. :2457693	Max. :2016	Max. :12.000	Max. :31.00
(Other):42462				

Table 5: Summary of Sunspot Numbers

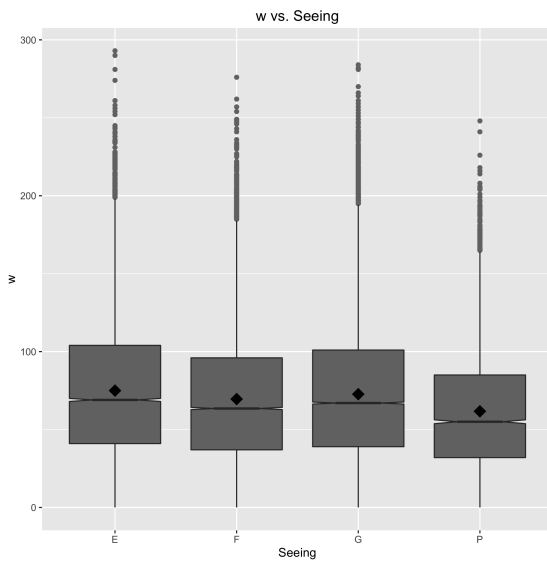
see	g	s	w	r	silso
E:10005	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :23195	n:36619
F:16636	1st Qu.: 3.000	1st Qu.: 10.00	1st Qu.: 38.00	3000F : 8920	y:17771
G:23294	Median : 4.000	Median : 22.00	Median : 65.00	2500E : 7082	
P: 4455	Mean : 4.386	Mean : 27.39	Mean : 71.25	3500G : 4228	
	3rd Qu.: 6.000	3rd Qu.: 39.00	3rd Qu.: 99.00	1000B : 3966	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3021	
				(Other): 3978	



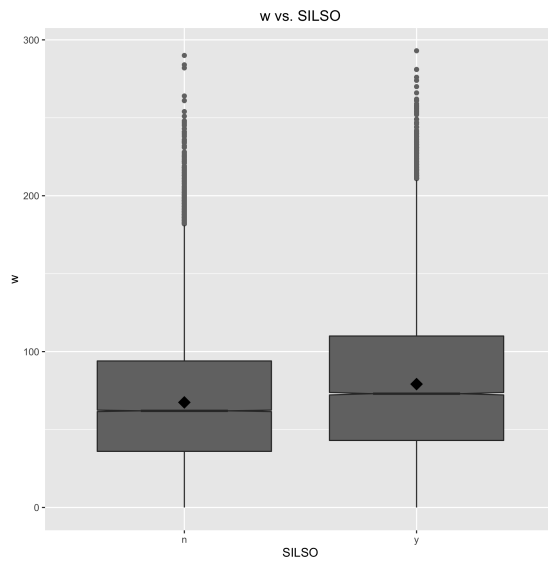
(a) Observed sunspot count histogram.



(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:

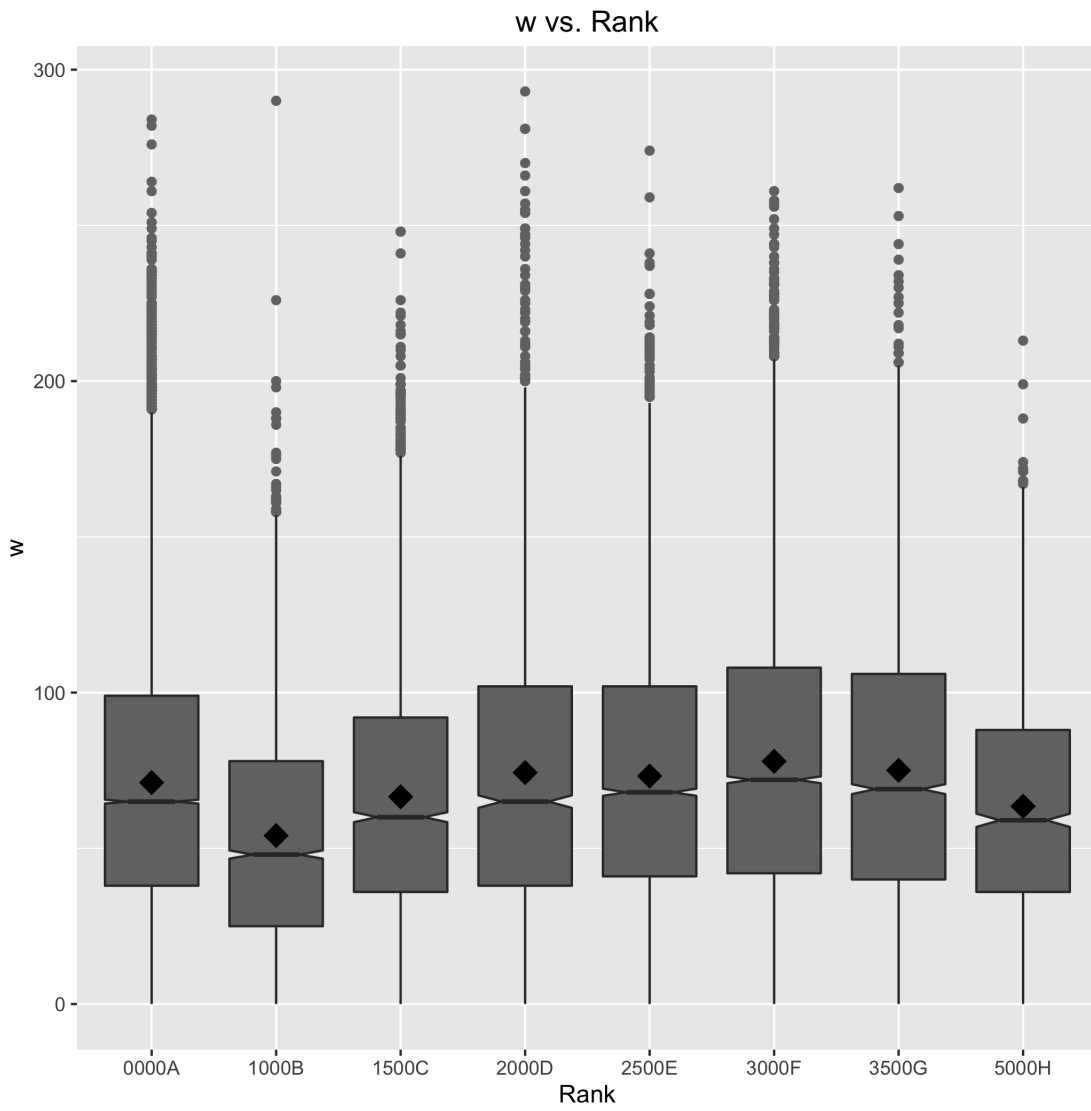


Figure 4: Box plot of sunspot count by rank.