

Monthly Report (00)

2016.11 Data Set

Saturday 17th December, 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.

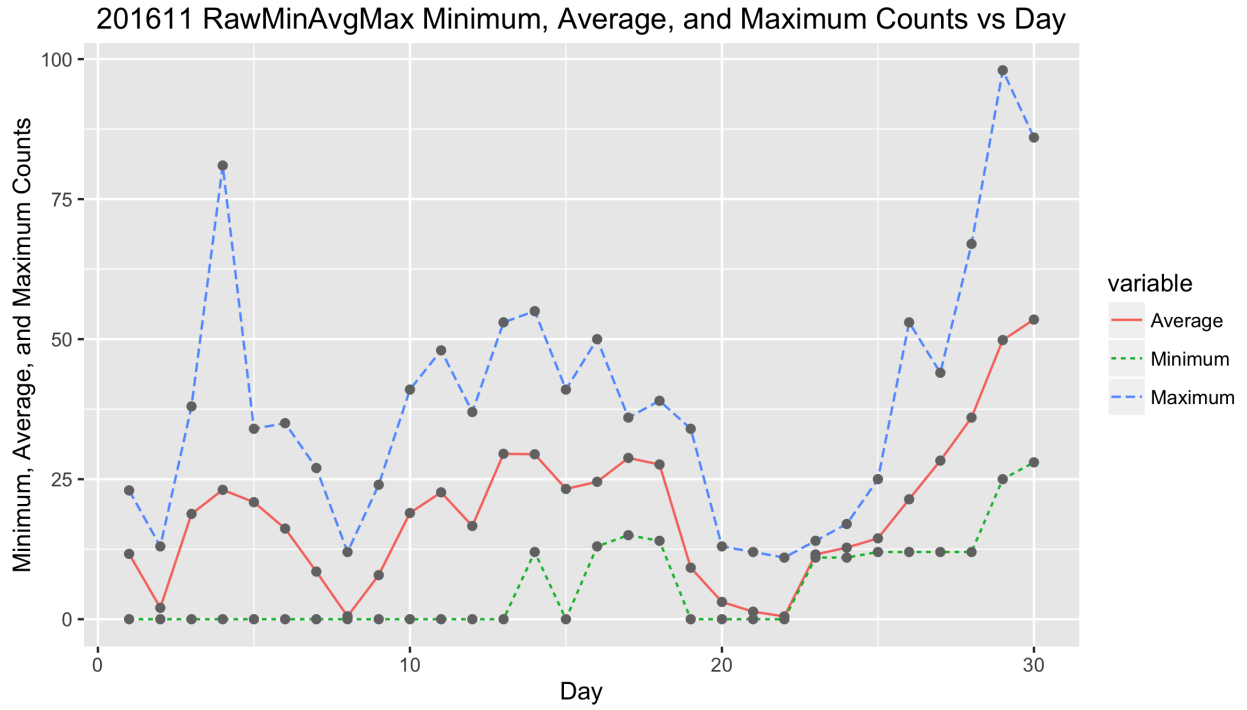


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201611 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	31.0000	0.0000	11.6667	23.0000
2.0000	32.0000	0.0000	2.0357	13.0000
3.0000	31.0000	0.0000	18.8000	38.0000
4.0000	24.0000	0.0000	23.0952	81.0000
5.0000	35.0000	0.0000	20.8966	34.0000
6.0000	32.0000	0.0000	16.1786	35.0000
7.0000	29.0000	0.0000	8.5000	27.0000
8.0000	24.0000	0.0000	0.5217	12.0000
9.0000	25.0000	0.0000	7.8750	24.0000
10.0000	28.0000	0.0000	18.9615	41.0000
11.0000	29.0000	0.0000	22.6538	48.0000
12.0000	34.0000	0.0000	16.6429	37.0000
13.0000	34.0000	0.0000	29.5357	53.0000
14.0000	23.0000	12.0000	29.4500	55.0000
15.0000	30.0000	0.0000	23.2593	41.0000
16.0000	36.0000	13.0000	24.5312	50.0000
17.0000	33.0000	15.0000	28.7931	36.0000
18.0000	27.0000	14.0000	27.6250	39.0000
19.0000	30.0000	0.0000	9.2000	34.0000
20.0000	26.0000	0.0000	3.0870	13.0000
21.0000	27.0000	0.0000	1.3462	12.0000
22.0000	23.0000	0.0000	0.5000	11.0000
23.0000	25.0000	11.0000	11.5652	14.0000
24.0000	28.0000	11.0000	12.7778	17.0000
25.0000	26.0000	12.0000	14.4400	25.0000
26.0000	24.0000	12.0000	21.4091	53.0000
27.0000	28.0000	12.0000	28.3182	44.0000
28.0000	24.0000	12.0000	36.0000	67.0000
29.0000	27.0000	25.0000	49.8333	98.0000
30.0000	26.0000	28.0000	53.5000	86.0000

3 Error Tables

Data are for the month of November 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	23.7956	23.2529	24.3382	8.4000	8.7000
2010.06	18.2224	17.7369	18.7079	11.0000	13.6000
2010.07	20.5031	20.0508	20.9554	15.2000	16.1000
2010.08	20.2150	19.7237	20.7064	18.3000	19.6000
2010.09	23.8013	23.2917	24.3109	22.8000	25.2000
2010.10	22.6013	22.1141	23.0885	21.0000	23.5000
2010.11	23.3312	22.8053	23.8571	20.9000	21.6000
2010.12	23.4222	22.7482	24.0962	13.9000	14.5000
2011.01	75.5241	73.8275	77.2208	17.7000	18.7000
2011.02	66.0646	64.5771	67.5521	29.1000	29.6000
2011.03	71.7379	70.2563	73.2194	48.0000	55.8000
2011.04	77.2330	75.5814	78.8846	47.3000	54.4000
2011.05	80.3968	78.7935	82.0000	37.3000	41.5000
2011.06	64.8693	63.5054	66.2332	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	71.1321	69.5682	72.6959	41.5000	43.8000
2011.08	73.4764	72.0372	74.9156	42.4000	50.5000
2011.09	83.7169	82.6157	84.8180	73.8000	78.0000
2011.10	78.9345	77.5791	80.2899	78.9000	88.0000
2011.11	80.1561	78.4387	81.8736	84.6000	96.7000
2011.12	78.1323	76.4171	79.8476	65.8000	73.0000
2012.01	77.5677	76.0285	79.1069	55.8000	58.2000
2012.02	65.6342	64.2401	67.0283	29.2000	33.1000
2012.03	74.1915	72.8638	75.5192	53.1000	64.1000
2012.04	76.9152	74.6490	79.1815	51.4000	55.2000
2012.05	84.3289	82.8550	85.8028	61.8000	69.0000
2012.06	67.8761	66.6706	69.0816	59.7000	64.5000
2012.07	75.3779	74.1073	76.6485	64.2000	51.3000
2012.08	74.1922	72.9513	75.4330	57.7000	63.1000
2012.09	84.5011	83.0560	85.9461	57.7000	61.5000
2012.10	81.1962	79.6617	82.7306	48.3000	53.3000
2012.11	83.5186	81.8437	85.1935	56.7000	61.4000
2012.12	79.1214	77.4534	80.7893	37.4000	40.8000
2013.01	87.5071	85.8659	89.1483	63.8000	62.9000
2013.02	75.7633	74.2858	77.2408	37.8000	38.0000
2013.03	81.5060	79.9719	83.0401	50.6000	57.9000
2013.04	89.6477	88.1385	91.1569	70.6000	72.4000
2013.05	92.1359	90.5459	93.7259	77.4000	78.7000
2013.06	74.7773	73.4458	76.1088	51.0000	52.5000
2013.07	80.9861	79.7269	82.2452	57.0000	57.0000
2013.08	81.8310	80.5535	83.1085	60.0000	66.0000
2013.09	92.3348	90.7395	93.9302	34.6000	36.9000
2013.10	87.2329	85.6828	88.7830	74.5000	85.6000
2013.11	89.8201	87.9528	91.6874	73.9000	77.6000
2013.12	87.2753	85.5199	89.0307	77.8000	90.3000
2014.01	104.0254	101.8385	106.2124	77.4000	82.0000
2014.02	90.0409	88.3292	91.7525	93.9000	102.8000
2014.03	100.2600	98.5668	101.9531	80.9000	92.2000
2014.04	109.0028	107.1510	110.8545	76.9000	84.7000
2014.05	111.5872	109.8022	113.3723	72.3000	75.2000
2014.06	90.5548	89.0858	92.0238	67.2000	71.0000
2014.07	99.4410	97.8191	101.0629	72.5000	72.5000
2014.08	99.9383	98.4384	101.4381	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	113.9546	112.1192	115.7901	83.2000	87.6000
2014.10	107.4353	105.6355	109.2350	59.5000	60.6000
2014.11	111.1794	109.0815	113.2773	65.8000	71.1000
2014.12	105.5038	103.2223	107.7853	75.8000	78.0000
2015.01	63.4325	62.2139	64.6510	65.9000	67.0000
2015.02	54.9049	53.6266	56.1832	42.4000	44.8000
2015.03	59.9718	58.8735	61.0701	38.0000	38.4000
2015.04	65.9873	64.8180	67.1567	49.0000	54.4000
2015.05	67.1429	66.0669	68.2189	56.3000	58.8000
2015.06	54.9548	54.0210	55.8885	50.2000	68.3000
2015.07	59.1346	58.1270	60.1422	47.9000	65.8000
2015.08	60.8982	59.9029	61.8935	39.5000	57.2000
2015.09	69.1493	68.0249	70.2738	49.2000	72.1000
2015.10	65.1512	64.0396	66.2629	39.3000	48.3000
2015.11	67.9173	66.9971	68.8375	39.6000	55.9000
2015.12	63.9717	62.6808	65.2626	36.4000	44.8000
2016.01	36.9039	36.2493	37.5585	33.7000	43.3000
2016.02	31.2690	30.6495	31.8886	38.3000	46.8000
2016.03	34.0340	33.4105	34.6576	30.5000	38.9000
2016.04	36.8188	36.1662	37.4714	26.6000	30.9000
2016.05	38.4882	37.8300	39.1464	33.7000	48.4000
2016.06	31.1344	30.6412	31.6276	13.1000	19.5000
2016.07	34.2227	33.6946	34.7507	21.2000	27.5000
2016.08	34.8767	34.3024	35.4510	33.0000	47.9000
2016.09	39.3691	38.7161	40.0221	27.7000	37.1000
2016.10	37.3543	36.7074	38.0013	22.7000	31.7000
2016.11	38.5212	37.8576	39.1849	14.0000	22.2000

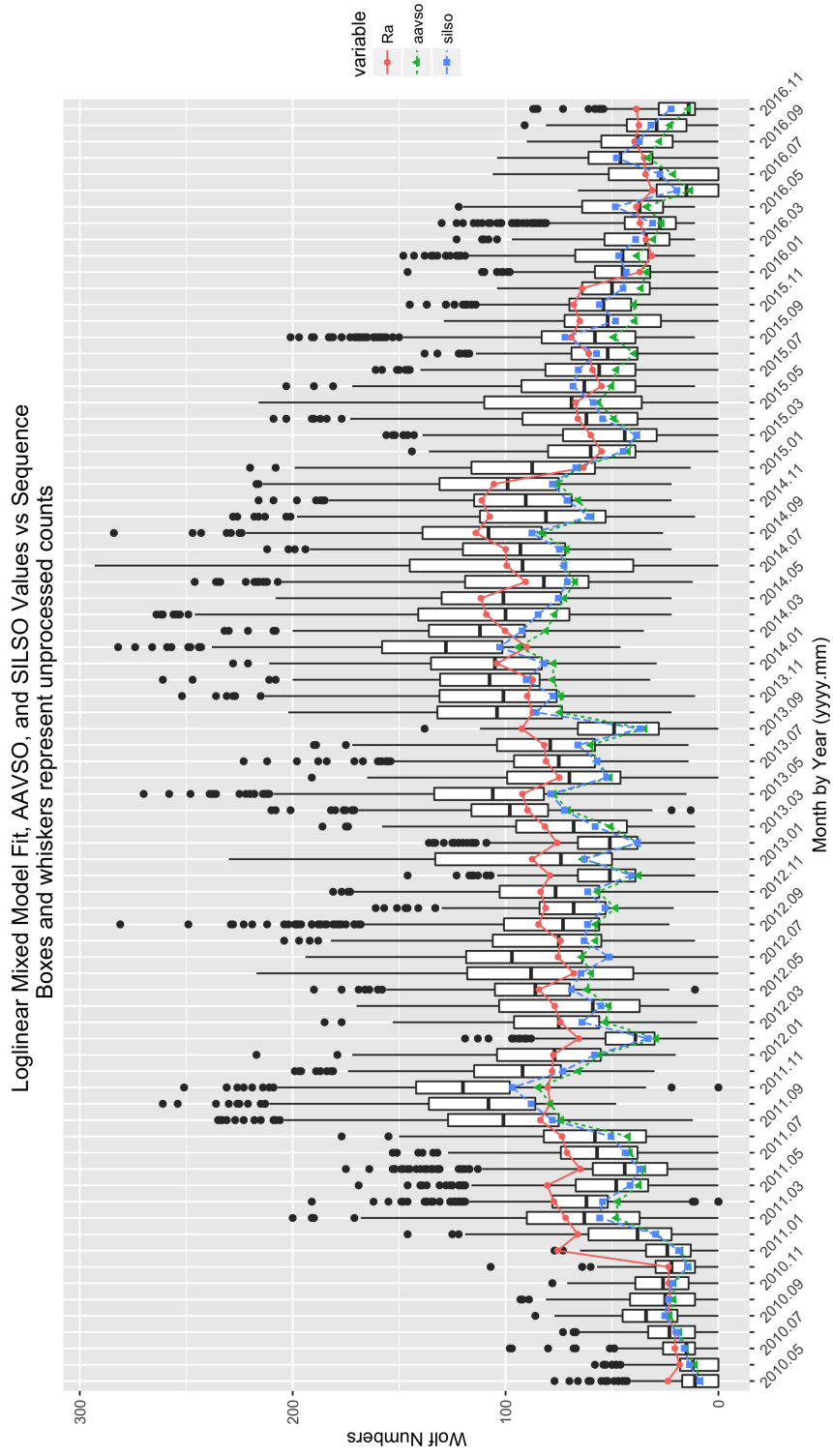


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201611 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.2108	0.0443	72.3997	0.0000
seeF	-0.1888	0.0072	-26.2081	0.0000
seeG	-0.1025	0.0063	-16.3681	0.0000
seeP	-0.2962	0.0106	-28.0071	0.0000
r1000B	-0.0561	0.0828	-0.6774	0.4981
r1500C	0.0357	0.1266	0.2823	0.7777
r2000D	0.0773	0.1544	0.5005	0.6168
r2500E	0.0000	0.1048	0.0004	0.9996
r3000F	0.0717	0.1021	0.7025	0.4823
r3500G	0.1216	0.1528	0.7955	0.4263
r5000H	-0.1073	0.2114	-0.5075	0.6118
silsoy	0.1217	0.0736	1.6536	0.0982
year2011	1.2072	0.0153	78.7962	0.0000
year2012	1.2239	0.0153	80.1464	0.0000
year2013	1.3203	0.0152	86.7054	0.0000
year2014	1.5095	0.0151	99.7939	0.0000
year2015	1.0114	0.0155	65.1734	0.0000
year2016	0.4486	0.0166	27.0434	0.0000
mon2	-0.1536	0.0120	-12.7572	0.0000
mon3	-0.0648	0.0111	-5.8548	0.0000
mon4	0.0217	0.0111	1.9452	0.0518
mon5	0.0484	0.0105	4.5910	0.0000
mon6	-0.1698	0.0111	-15.2369	0.0000
mon7	-0.0814	0.0107	-7.6112	0.0000
mon8	-0.0653	0.0105	-6.1992	0.0000
mon9	0.0662	0.0102	6.5205	0.0000
mon10	0.0132	0.0107	1.2349	0.2169
mon11	0.0516	0.0110	4.6972	0.0000
mon12	0.0099	0.0118	0.8431	0.3992

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.

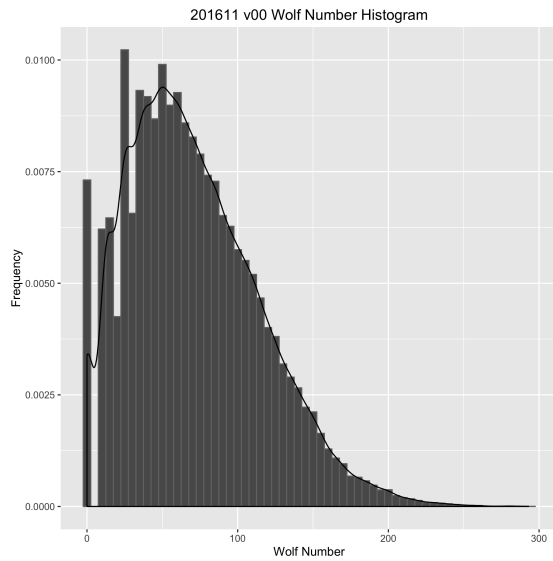
5 Supporting Information

Table 4: 201611 Summary of Sunspot Numbers

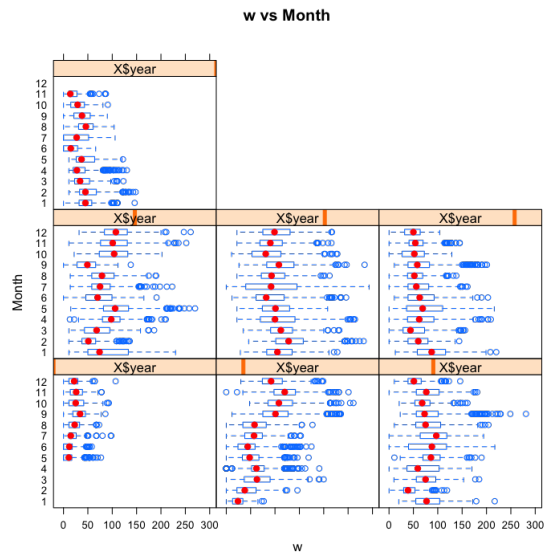
obs	jd	year	mon	day
ARAG : 2347	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 2156	1st Qu.:2455994	1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 2129	Median :2456536	Median :2013	Median : 7.000	Median :16.00
BROB : 1906	Mean :2456232	Mean :2013	Mean : 6.722	Mean :15.72
HOWR : 1774	3rd Qu.:2457136	3rd Qu.:2015	3rd Qu.: 9.000	3rd Qu.:23.00
DUBF : 1750	Max. :2457723	Max. :2016	Max. :12.000	Max. :31.00
(Other):42877				

Table 5: Summary of Sunspot Numbers

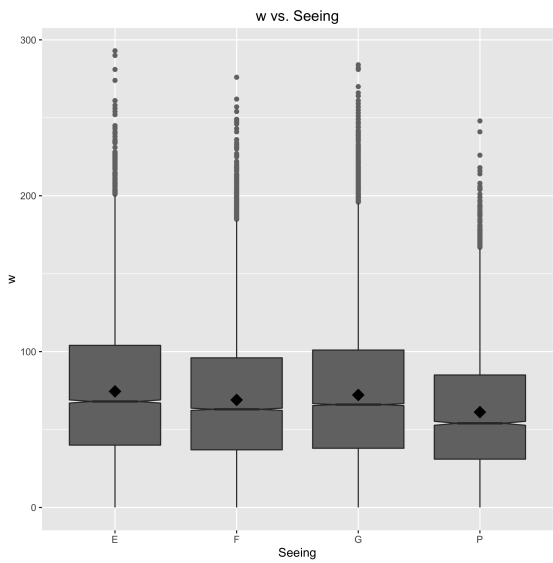
see	g	s	w	r	silso
E:10096	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :23401	n:36982
F:16822	1st Qu.: 2.000	1st Qu.: 10.00	1st Qu.: 37.00	3000F : 9045	y:17957
G:23512	Median : 4.000	Median : 21.00	Median : 65.00	2500E : 7163	
P: 4509	Mean : 4.356	Mean : 27.17	Mean : 70.73	3500G : 4285	
	3rd Qu.: 6.000	3rd Qu.: 39.00	3rd Qu.: 99.00	1000B : 3997	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 3021	
				(Other): 4027	



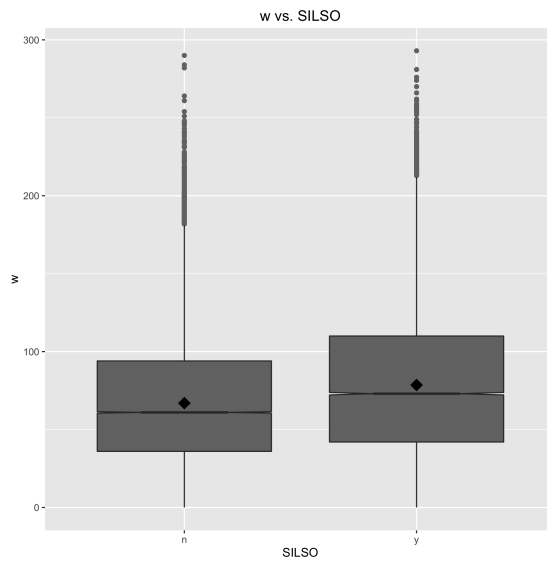
(a) Observed sunspot count histogram.



(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:

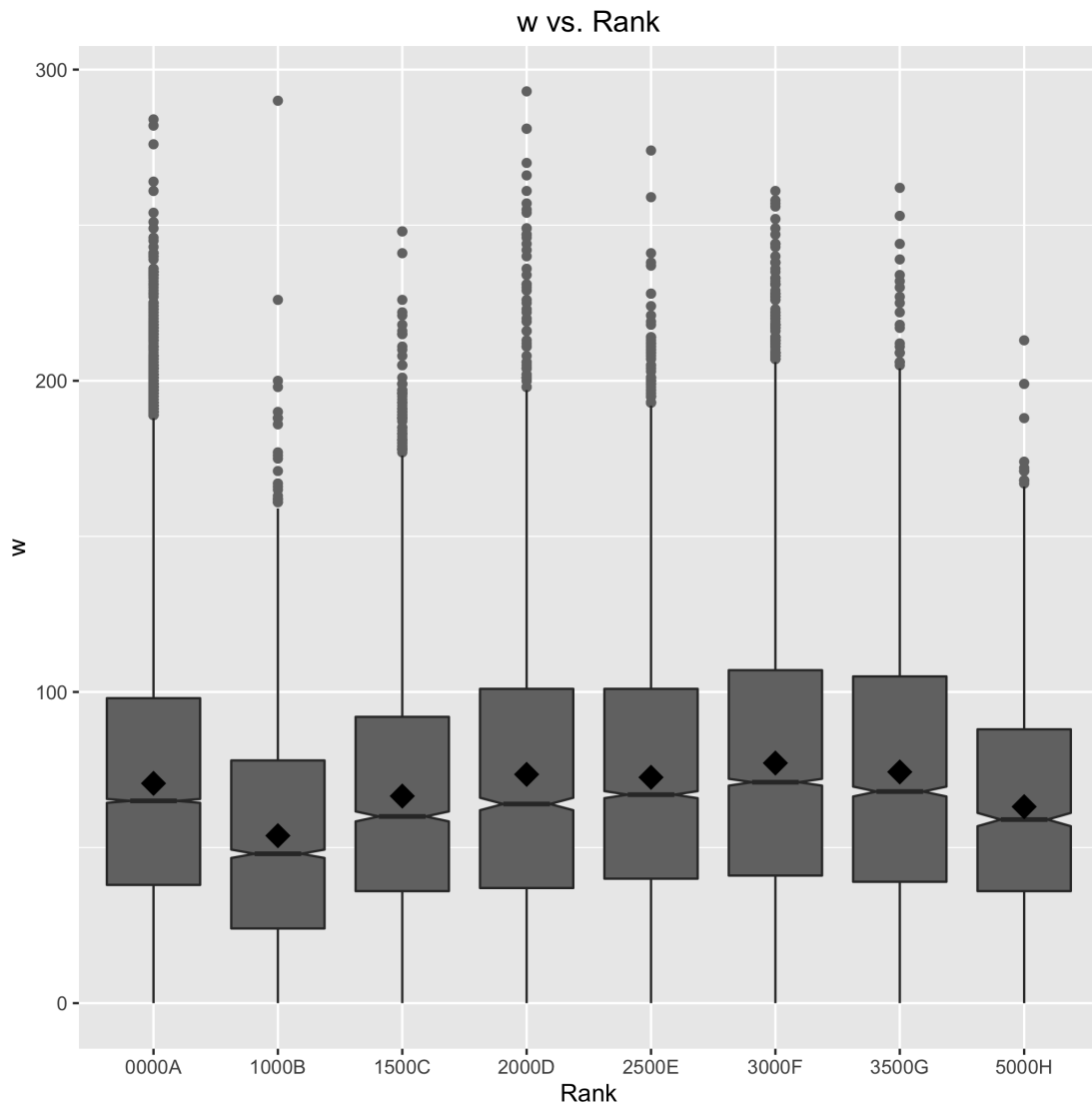


Figure 4: Box plot of sunspot count by rank.