Monthly Report (00)

2015.11 Data Set

Wednesday 16th December, 2015

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.

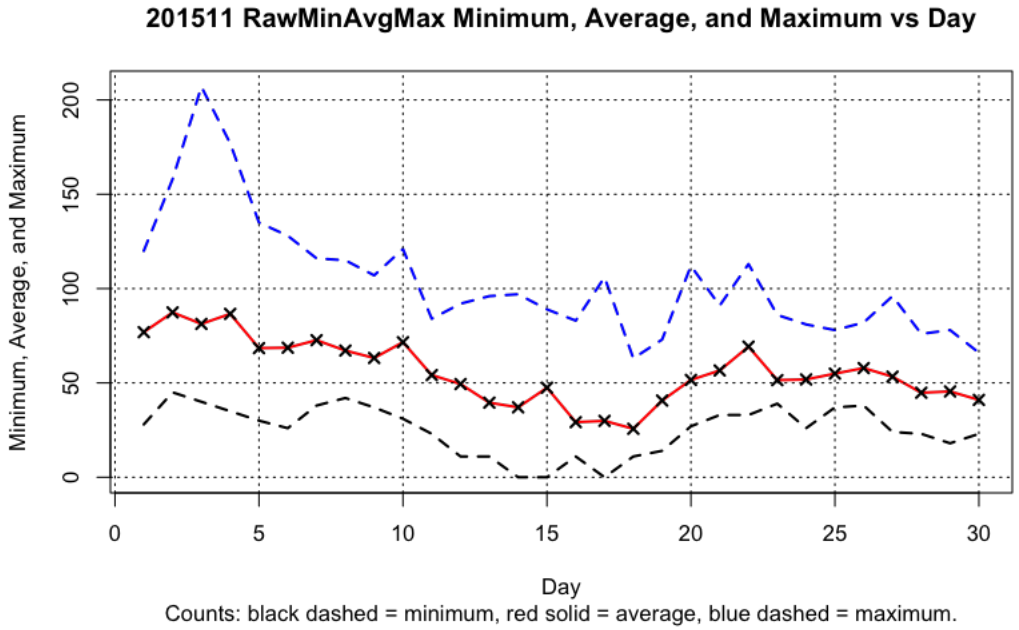


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201511 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	33.0000	28.0000	76.9167	120.0000
2.0000	28.0000	45.0000	87.3750	158.0000
3.0000	34.0000	40.0000	81.2759	207.0000
4.0000	27.0000	35.0000	86.5926	177.0000
5.0000	28.0000	30.0000	68.4444	135.0000
6.0000	24.0000	26.0000	68.6667	128.0000
7.0000	24.0000	38.0000	72.6522	116.0000
8.0000	33.0000	42.0000	67.1429	115.0000
9.0000	35.0000	37.0000	63.2143	107.0000
10.0000	22.0000	31.0000	71.5789	121.0000
11.0000	22.0000	23.0000	54.1500	84.0000
12.0000	29.0000	11.0000	49.4074	92.0000
13.0000	24.0000	11.0000	39.5217	96.0000
14.0000	20.0000	0.0000	37.0526	97.0000
15.0000	26.0000	0.0000	47.4545	89.0000
16.0000	31.0000	11.0000	29.1923	83.0000
17.0000	29.0000	0.0000	29.8846	106.0000
18.0000	29.0000	11.0000	25.6923	63.0000
19.0000	28.0000	14.0000	40.6538	73.0000
20.0000	29.0000	27.0000	51.5926	112.0000
21.0000	32.0000	33.0000	56.5862	91.0000
22.0000	30.0000	33.0000	69.2400	113.0000
23.0000	32.0000	39.0000	51.4828	86.0000
24.0000	28.0000	26.0000	51.8400	81.0000
25.0000	31.0000	37.0000	54.9259	78.0000
26.0000	24.0000	38.0000	57.8636	82.0000
27.0000	27.0000	24.0000	53.3182	96.0000
28.0000	24.0000	23.0000	44.8000	76.0000
29.0000	25.0000	18.0000	45.4762	78.0000
30.0000	24.0000	23.0000	41.0000	66.0000

3 Error Tables

Data are for the month of November 2015. 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
2010.05	23.2526	22.7231	23.7820	8.4000	8.7000
2010.06	18.7138	18.2151	19.2126	11.0000	13.6000
2010.07	20.5259	20.0750	20.9768	15.2000	16.1000
2010.08	19.5744	19.0990	20.0498	18.3000	19.6000
2010.09	23.5779	23.0747	24.0811	22.8000	25.2000
2010.10	22.6125	22.1246	23.1004	21.0000	23.5000
2010.11	23.8895	23.3502	24.4288	20.9000	21.6000
2010.12	24.0582	23.3682	24.7483	13.9000	14.5000
2011.01	73.5982	71.9471	75.2493	17.7000	18.7000
2011.02	62.1128	60.7122	63.5135	29.1000	29.6000
2011.03	70.7767	69.3195	72.2340	48.0000	55.8000
2011.04	77.4921	75.8285	79.1558	47.3000	54.4000
2011.05	79.0538	77.4644	80.6432	37.3000	41.5000
2011.06	67.0903	65.6760	68.5045	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	71.7058	70.1302	73.2813	41.5000	43.8000
2011.08	71.6819	70.2771	73.0868	42.4000	50.5000
2011.09	83.6672	82.5763	84.7581	73.8000	78.0000
2011.10	79.7966	78.4373	81.1560	78.9000	88.0000
2011.11	82.8893	81.1338	84.6449	84.6000	96.7000
2011.12	80.9748	79.2179	82.7318	65.8000	73.0000
2012.01	75.7183	74.2224	77.2141	55.8000	58.2000
2012.02	61.7958	60.4955	63.0961	29.2000	33.1000
2012.03	73.3526	72.0484	74.6568	53.1000	64.1000
2012.04	77.3231	75.0893	79.5569	51.4000	55.2000
2012.05	83.1636	81.7166	84.6107	61.8000	69.0000
2012.06	70.3195	69.0738	71.5651	59.7000	64.5000
2012.07	76.1563	74.8821	77.4304	64.2000	51.3000
2012.08	72.5778	71.3697	73.7859	57.7000	63.1000
2012.09	84.5504	83.1118	85.9891	57.7000	61.5000
2012.10	82.1091	80.5625	83.6557	48.3000	53.3000
2012.11	86.3989	84.6727	88.1251	56.7000	61.4000
2012.12	82.0109	80.2813	83.7405	37.4000	40.8000
2013.01	85.4085	83.8031	87.0140	63.8000	62.9000
2013.02	71.2151	69.8299	72.6004	37.8000	38.0000
2013.03	80.4083	78.9013	81.9152	50.6000	57.9000
2013.04	90.0681	88.5530	91.5832	70.6000	72.4000
2013.05	90.8382	89.2749	92.4014	77.4000	78.7000
2013.06	77.4394	76.0624	78.8163	51.0000	52.5000
2013.07	81.7312	80.4581	83.0042	57.0000	57.0000
2013.08	79.9476	78.6988	81.1963	60.0000	66.0000
2013.09	92.3216	90.7291	93.9140	34.6000	36.9000
2013.10	88.2288	86.6654	89.7922	74.5000	85.6000
2013.11	92.9317	91.0024	94.8610	73.9000	77.6000
2013.12	90.5556	88.7337	92.3776	77.8000	90.3000
2014.01	101.5929	99.4613	103.7245	77.4000	82.0000
2014.02	84.8487	83.2419	86.4556	93.9000	102.8000
2014.03	99.0294	97.3615	100.6973	80.9000	92.2000
2014.04	109.5179	107.6587	111.3771	76.9000	84.7000
2014.05	110.0849	108.3279	111.8419	72.3000	75.2000
2014.06	93.8780	92.3529	95.4031	67.2000	71.0000
2014.07	100.5007	98.8681	102.1333	72.5000	72.5000
2014.08	97.7830	96.3101	99.2558	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	113.9167	112.0973	115.7361	83.2000	87.6000
2014.10	108.6451	106.8283	110.4619	59.5000	60.6000
2014.11	115.0130	112.8354	117.1906	65.8000	71.1000
2014.12	109.3496	107.0048	111.6944	75.8000	78.0000
2015.01	62.7068	61.5052	63.9083	65.9000	67.0000
2015.02	52.3345	51.1330	53.5361	42.4000	44.8000
2015.03	60.0260	58.9354	61.1167	38.0000	38.4000
2015.04	67.2257	66.0428	68.4086	49.0000	54.4000
2015.05	67.2906	66.2091	68.3722	56.3000	58.8000
2015.06	57.7577	56.7827	58.7326	50.2000	68.3000
2015.07	60.4908	59.4658	61.5158	47.9000	66.4000
2015.08	60.2386	59.2572	61.2200	39.5000	64.6000
2015.09	70.0157	68.8815	71.1499	49.2000	78.1000
2015.10	66.7111	65.5763	67.8458	39.3000	61.7000
2015.11	71.1680	70.2068	72.1292	39.6000	63.2000

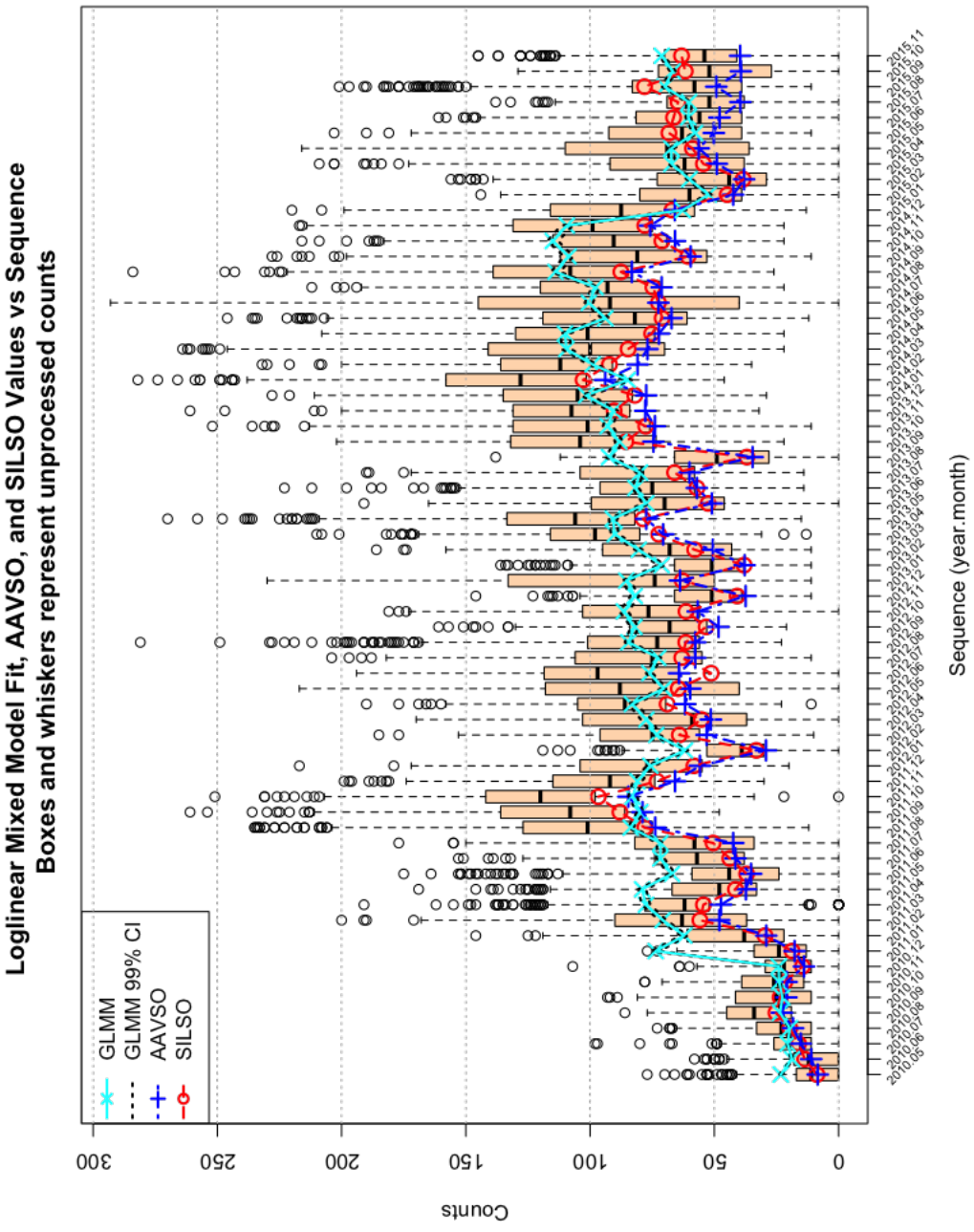


Figure 2: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

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: 201511 Parameter Estimates

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.1796	0.0446	71.3545	0.0000
seeF	-0.1842	0.0075	-24.6772	0.0000
seeG	-0.0968	0.0065	-14.8920	0.0000
seeP	-0.2940	0.0110	-26.7740	0.0000
r1000B	-0.0626	0.0830	-0.7540	0.4509
r1500C	0.0274	0.1269	0.2162	0.8288
r2000D	0.0762	0.1548	0.4920	0.6227
r2500E	-0.0006	0.1051	-0.0054	0.9957
r3000F	0.0604	0.1024	0.5898	0.5553
r3500G	0.1177	0.1532	0.7685	0.4422
r5000H	-0.1149	0.2119	-0.5422	0.5877
silsoy	0.1184	0.0738	1.6042	0.1087
year2011	1.2158	0.0153	79.4287	0.0000
year2012	1.2329	0.0153	80.7985	0.0000
year2013	1.3285	0.0152	87.3153	0.0000
year2014	1.5185	0.0151	100.4481	0.0000
year2015	1.0336	0.0156	66.3207	0.0000
mon2	-0.1900	0.0128	-14.8579	0.0000
mon3	-0.0533	0.0116	-4.5936	0.0000
mon4	0.0512	0.0117	4.3958	0.0000
mon5	0.0588	0.0110	5.3301	0.0000
mon6	-0.1097	0.0116	-9.4868	0.0000
mon7	-0.0472	0.0112	-4.2278	0.0000
mon8	-0.0640	0.0110	-5.7933	0.0000
mon9	0.0905	0.0106	8.5432	0.0000
mon10	0.0490	0.0112	4.3889	0.0000
mon11	0.1099	0.0114	9.6702	0.0000
mon12	0.0706	0.0125	5.6627	0.0000

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.

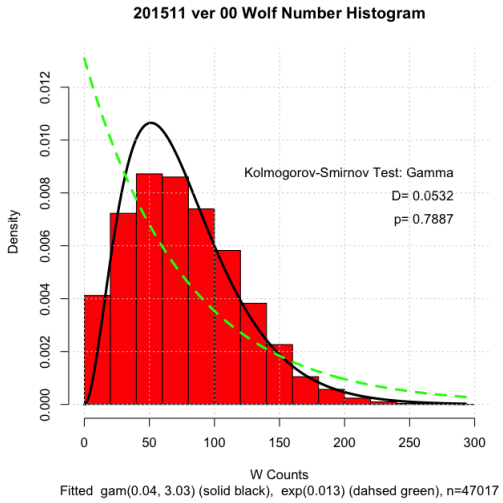
5 Supporting Information

Table 4: 201511 Summary of Sunspot Numbers

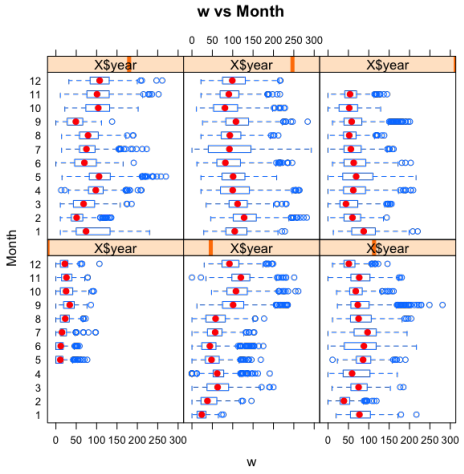
obs	jd	year	mon	day
ARAG : 1967	Min. :1721096	Min. :2010	Min. : 1.00	Min. : 1.00
CHAG : 1793	1st Qu.:2455891	1st Qu.:2011	1st Qu.: 4.00	1st Qu.: 8.00
BRAB : 1759	Median :2456389	Median :2013	Median : 7.00	Median :16.00
BROB : 1585	Mean :2455995	Mean :2013	Mean : 6.71	Mean :15.73
DUBF : 1486	3rd Qu.:2456852	3rd Qu.:2014	3rd Qu.: 9.00	3rd Qu.:23.00
HOWR : 1445	Max. :2457357	Max. :2015	Max. :12.00	Max. :31.00
(Other):36455				

Table 5: Summary of Sunspot Numbers

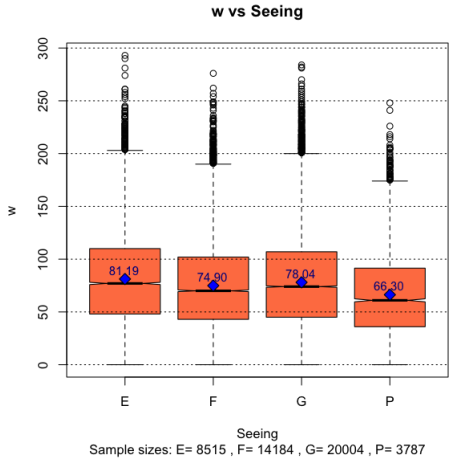
see	g	s	w	r	silso
E: 8515	Min. : 0.000	Min. : 0.00	Min. : 0.0	0000A :20037	n:31502
F:14184	1st Qu.: 3.000	1st Qu.: 12.00	1st Qu.: 44.0	3000F : 7458	y:14988
G:20004	Median : 5.000	Median : 24.00	Median : 72.0	2500E : 6004	
P: 3787	Mean : 4.698	Mean : 29.72	Mean : 76.7	3500G : 3552	
	3rd Qu.: 6.000	3rd Qu.: 42.00	3rd Qu.:105.0	1000B : 3385	
	Max. :18.000	Max. :204.00	Max. :293.0	1500C : 2779	
				(Other): 3275	



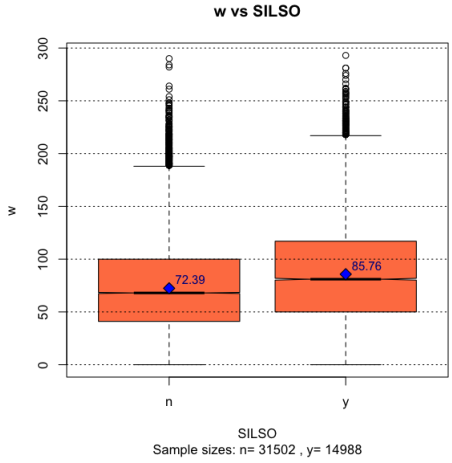
(a) Observed sunspot count histogram.



(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:

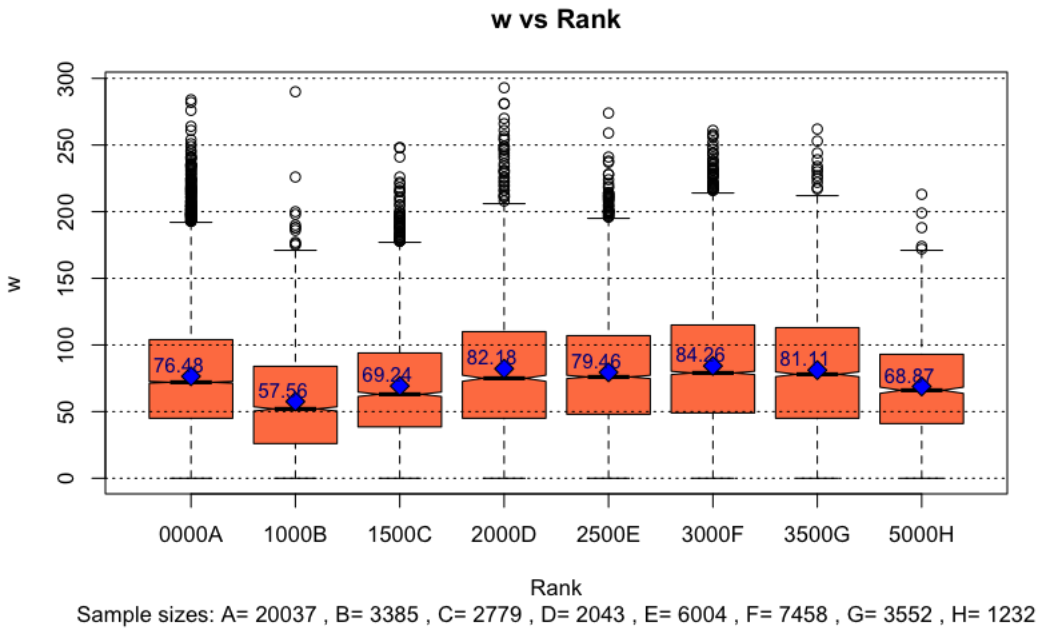


Figure 4: Box plot of sunspot count by rank.