Solar Bulletin

THE AMERICAN ASSOCIATION OF VARIABLE STAR OBSERVERS
SOLAR SECTION

Rodney Howe, Kristine Larsen, Co-Chairs
c/o AAVSO, 185 Alewife Brook Parkway, # 410
Cambridge, MA 02138 USA

Web: http://www.aavso.org/solar-bulletin
Email: solar@aavso.org
ISSN 0271-8480

Volume 78 Number 8 August 2022

The Solar Bulletin of the AAVSO is a summary of each month's solar activity recorded by visual solar observers' counts of group and sunspots, and the VLF radio recordings of SID Events in the ionosphere. The sudden ionospheric disturbance report is in Section 2. The relative sunspot numbers are in Section 3. Section 4 has endnotes.

1 Sunspot Group Numbers Compared

Submitted by Jamie Riggs, Ph.D.

A common question in determining monthly Sunspot numbers is how similar are the counts across observers after accounting for seeing, equipment used, filter type, etc. The object is to account for group count consistency across observers. A first step, an exploratory data analysis (EDA) component, is to visually compare the group counts of selected observers who post Sunspot counts to the AAVSO Solar Section database [AAVSO, 2022]. These selected observers submitted counts each month during the years spanning January 2000 to August 2022, which defines the study period.

Data for observers contributing Sunspot numbers over the study period were subset to a manageable dozen observers. (More observers will be included as this study progresses.) The data for these observers were screened to assure there is no missing data throughout the 22-year period. Smoothing is employed to impute missing group counts as needed. The EDA and resulting data conditioning prepared these data for modeling using the DeepAR recurrent neural network methodology implemented by Amazon, Inc. for multiple concurrent time series data.

A recurrent neural network (RNN) is a neural network in which node layers follow a time-ordered sequence. RNN nodes act as "memory" for forward feeding state changes. This makes them viable for modeling time series and thus allows for forecasting. See, for example, `https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks` for more details.

The layout of the nodes of a neural network have connections that are one way, namely, the update rule of propagation through time is forward. Time series have inherent autocorrelation, i.e., past realizations of the time series affect the state of the current realization. The neural network needs a feedback layout in addition to the feedforward layout structure. The feedforward/feedback structure is called a recurrent neural network. There are many different ways in which feedback can occur; for example, back from the output to the hidden layer nodes, or feedback from among the nodes. See Figure 1 for one graphical representation of the feedback and feedforward structure of an RNN.

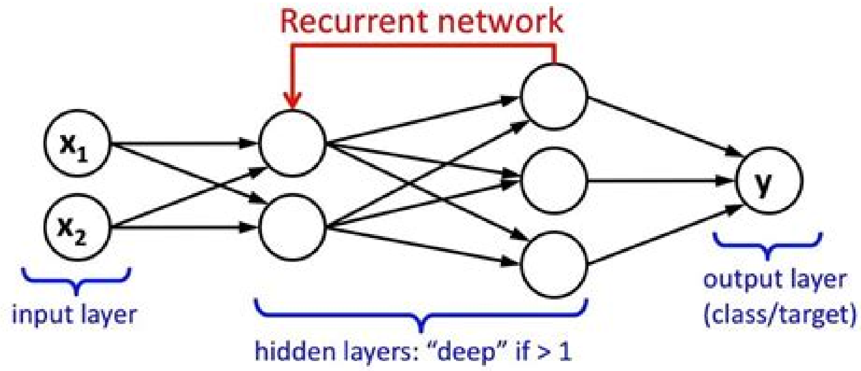


Figure 1: `https://www.researchgate.net/figure/`. The inputs x_1 and x_2 denote two Solar Section observers' group counts. The y output is the estimated group count for one time increment. The hidden layers represent the feedforward and feedback system of the RNN.

Recurrent networks are used to learn sequential inputs such as in speech recognition or in machine translations from one language to another. The training set consists of time sequences $\{\boldsymbol{x}(t), \boldsymbol{y}(t)\}$ of inputs and outputs. The network is trained on the sequences and learns to predict the outputs. The inputs and the outputs depend on and uses a discrete-time update rule. Rather than just one output as in neural networks, recurrent neural networks have separate outputs as a function of each value of t.

To train recurrent networks with time-dependent inputs, the RNN uses back-propagation through time. The network is unfolded in time to implement the feedback. The penalty for this unfolding is large networks, with as many nodes as there are time increments.

The Amazon DeepAR forecasting algorithm is a supervised learning algorithm for forecasting scalar (one-dimensional) time series using recurrent neural networks (RNN). DeepAR uses autoregeressive (time-dependent) RNNs to model multivariate time series with identical date and time stamps. Examples include market indices such as the DOW and S&P 500, a suite of stocks, shipping logistics models, and multiple store and store department daily sales. DeepAR can outperform classical time series models or smoother-based time series models when there are many (can be hundreds) time series. As is the case with all time series models, DeepAR also generates forecasts. Details on Amazon DeepAR may be found at `https://docs.aws.amazon. com/sagemaker/latest/dg/deepar.html`.

The objective of using DeepAR on the 12 AAVSO observers' Sunspot group counts is to find forecasts for each observer based on each observer's group count history, including possible latent interplay among the observers. For this report, only a visual comparison is made. A statistical comparison is slated for an ensuing Solar Bulletin edition.

Figure 2 shows the individual historical group counts (black curves) with 48 months of forecasted (red curves) group counts, one plot per observer. Observers ARAG, BROB, BXZ, CKB, DUBF, FLET, HKY, KNJS, and RJV have visually similar group count histories and thereby forecasts as determined by the DeepAR algorithm. Observers BARH and CHAG have similar histories and forecasts. Observer SIMC is possibly unique among this group of 12 observers.

These 12 observers visually show consistency within two collections of observers with one observer seemingly unique relative to the other 11 observers. Visual representations are useful for guiding the next steps of the analysis. These next steps include comparing the 12 observer counts using multivariate time series statistical comparisons to test how, if at all, the observers histories and forecasts are the same. Once the statistical tests are implemented, a further step is to include

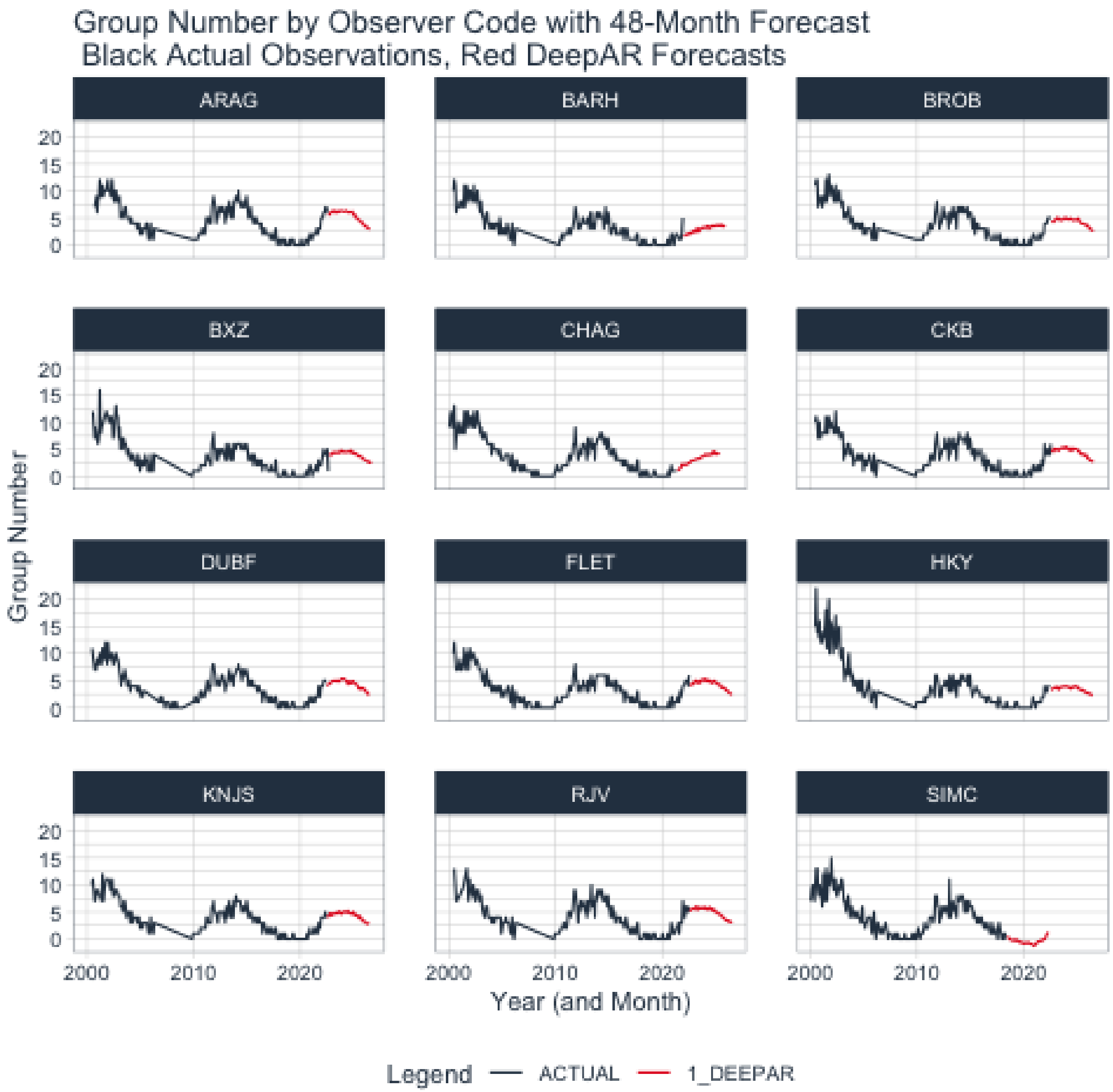


Figure 2: Twelve AAVSO observers' group counts. The black curves are the historical group counts (ACTUAL) from January 2000 to August 2022. The red curves (1_DEEPAR) are the observers' forecasts from September 2022 to August 2026.

as many of the Solar Section observers in the analysis as the submission data allow. Finally, is there a combination of the observer counts that lead to a useful set of forecasts of Sunspot groups? If so, we wish to examine the time series autocorrelation structure, and the across-observer correlation structure. These correlations will inform us as to whether there is a latent, unintended connection among the observers that may allow a reliable, accurate forecast of the group counts.

2 Sudden Ionospheric Disturbance (SID) Report

2.1 SID Records

August 2022 (Figure 3), there were a group of C- and M- class flares leading up to a M1.8 flare on the 27th of August recorded by Roberto Battaiola in Milan Italy.

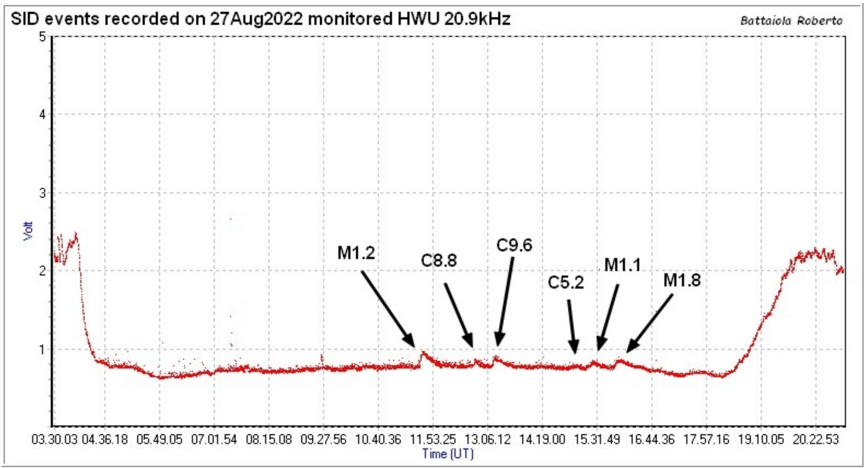


Figure 3: VLF recording on the 27th of August.

2.2 SID Observers

In August 2022, 18 AAVSO SID observers submitted VLF data as listed in Table 1.

Table 1: 202208 VLF Observers

Observer	Code	Stations
R Battaiola	A96	HWU
J Wallace	A97	NAA
L Loudet	A118	DHO
J Godet	A119	GBZ GQD ICV
B Terrill	A120	NWC
F Adamson	A122	NWC
G Perry	A126	DHO
J Karlovsky	A131	DHO NAA TBB
R Green	A134	NWC
R Mrllak	A136	GQD NSY
S Aguirre	A138	NPM NAA
G Silvis	A141	NAA NML NLK
K Menzies	A146	NAA
L Pina	A148	NAA NLK NML
J Wendler	A150	NAA
H Krumnow	A152	FTA GBZ HWU
J DeVries	A153	NLK
R Mazur	A155	NLK NML

Figure 4 depicts the importance rating of the solar events. The duration in minutes are -1: LT 19, 1: 19-25, 1+: 26-32, 2: 33-45, 2+: 46-85, 3: 86-125, and 3+: GT 125.

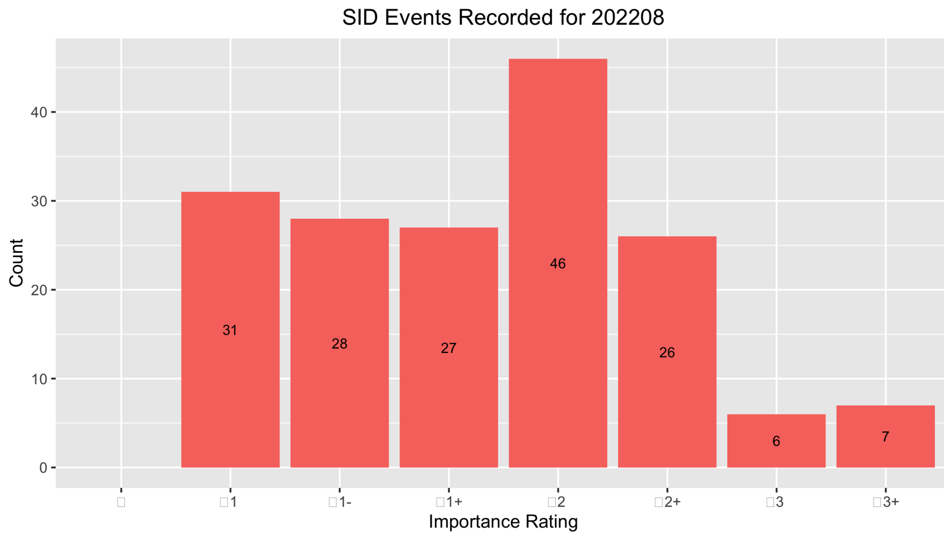


Figure 4: VLF SID Events.

2.3 Solar Flare Summary from GOES-16 Data

In August 2022, there were 346 GOES-16 XRA flares: 29 M-Class, 206 C-Class and 111 B-Class flares. Far more flaring this month than last (Figure 5).

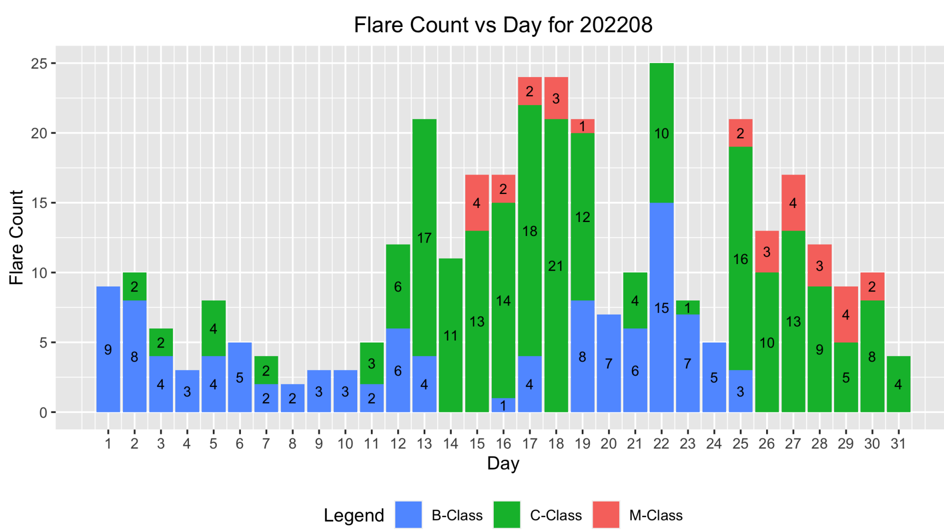


Figure 5: GOES-16 XRA (NOAA, 2022) flares.

3 Relative Sunspot Numbers (R_a)

Reporting monthly sunspot numbers consists of submitting an individual observer's daily counts for a specific month to the AAVSO Solar Section. These data are maintained in a Structured Query Language (SQL) database. The monthly data then are extracted for analysis. This section is the portion of the analysis concerned with both the raw and daily average counts for a particular month. Scrubbing and filtering the data assure error-free data are used to determine the monthly sunspot numbers.

3.1 Raw Sunspot Counts

The raw daily sunspot counts consist of submitted counts from all observers who provided data in August 2022. These counts are reported by the day of the month. 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 6.

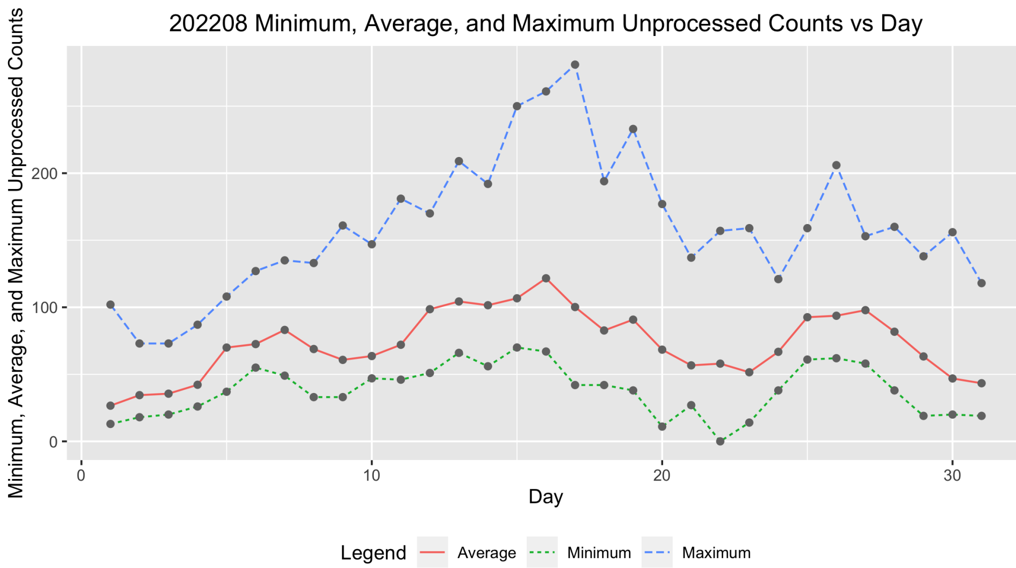


Figure 6: Raw Wolf number average, minimum and maximum by day of the month for all observers.

3.2 American Relative Sunspot Numbers

The relative sunspot numbers, R_a, contain the sunspot numbers after the submitted data are scrubbed and modeled by Shapley's method with k-factors (`https://adsabs.harvard.edu/full/1949PASP...61...13S`). The Shapley method is a statistical model that agglomerates variation due to random effects, such as observer group selection, and fixed effects, such as seeing condition. The raw Wolf averages and calculated R_a are seen in Figure 7, and Table 2 shows the Day of the observation (column 1), the Number of Observers recording that day (column 2), the Raw Wolf number (column 3), and the Shapley Correction (R_a) (column 4).

Table 2: 202208 American Relative Sunspot Numbers ($\mathrm{R_a}$).

Day	Number of Observers	Raw	R_a
1	47	29	24
2	45	36	31
3	43	35	31
4	43	43	39
5	46	71	64
6	47	75	65
7	49	84	74
8	51	71	59
9	45	62	53
10	43	65	56
11	46	73	63
12	51	99	86
13	43	109	92
14	49	101	87
15	49	107	91
16	40	125	101
17	34	106	88
18	37	93	77
19	44	99	80
20	47	72	57
21	41	59	49
22	38	57	47
23	42	51	44
24	45	68	60
25	41	94	82
26	41	93	78
27	44	100	85
28	42	83	70
29	45	62	52
30	41	48	40
31	40	41	36
Averages	43.8	74.5	63.3

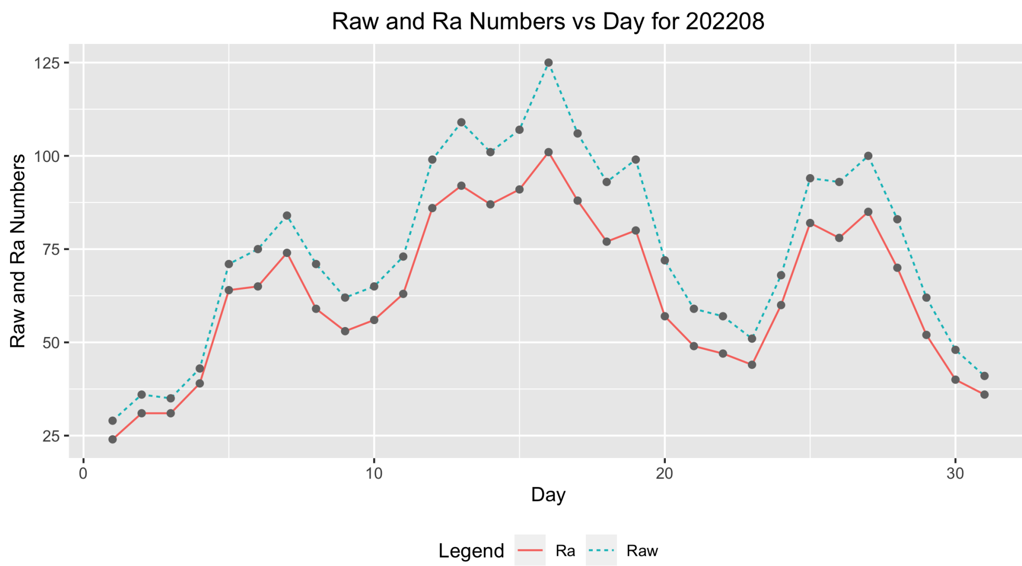


Figure 7: Raw Wolf average and R_a numbers by day of the month for all observers.

3.3 Sunspot Observers

Table 3 lists the Observer Code (column 1), the Number of Observations (column 2) submitted for August 2022, and the Observer Name (column 3). The final row gives the total number of observers who submitted sunspot counts (72), and total number of observations submitted (1359).

Table 3: 202208 Number of observations by observer.

Observer Code	Number of Observations	Observer Name
AAX	19	Alexandre Amorim
AJV	22	J. Alonso
ARAG	31	Gema Araujo
ASA	17	Salvador Aguirre
ATE	27	Teofilo Arranz Heras
BATR	11	Roberto Battaiola
BKL	16	John A. Blackwell
BMF	29	Michael Boschat
BMIG	30	Michel Besson
BROB	30	Robert Brown
BXZ	28	Jose Alberto Berdejo
BZX	26	A. Gonzalo Vargas
CANG	2	Andrew Corkill
CIOA	7	Ioannis Chouinavas
CKB	25	Brian Cudnik
CMOD	2	Mois Carlo
CNT	30	Dean Chantiles
CVJ	9	Jose Carvajal

Continued

Table 3: 202208 Number of observations by observer.

Observer Code	Number of Observations	Observer Name
DARB	12	Aritra Das
DJOB	12	Jorge del Rosario
DMIB	30	Michel Deconinck
DUBF	29	Franky Dubois
EHOA	11	Howard Eskildsen
ERB	27	Bob Eramia
FERA	26	Eric Fabrigat
FLET	23	Tom Fleming
GIGA	30	Igor Grageda Mendez
HALB	21	Brian Halls
HKY	24	Kim Hay
HMQ	6	Mark Harris
HOWR	27	Rodney Howe
HRUT	26	Timothy Hrutkay
IEWA	21	Ernest W. Iverson
ILUB	4	Luigi Iapichino
JDAC	9	David Jackson
JGE	9	Gerardo Jimenez Lopez
JSI	10	Simon Jenner
KAND	29	Kandilli Observatory
KAPJ	23	John Kaplan
KNJS	27	James & Shirley Knight
LEVM	9	Monty Leventhal
LJAE	12	Jay Lavender
LKR	11	Kristine Larsen
LRRA	9	Robert Little
MARC	6	Arnaud Mengus
MARE	16	Enrico Mariani
MCE	21	Etsuiku Mochizuki
MJAF	29	Juan Antonio Moreno Quesada
MJHA	29	John McCammon
MLL	14	Jay Miller
MMAY	31	Max Surlaroute
MMI	30	Michael Moeller
MSS	12	Sandy Mesics
MUDG	7	George Mudry
MWU	12	Walter Maluf
OAAA	25	Al Sadeem Astronomy Obs.
ONJ	16	John O'Neill
PLUD	28	Ludovic Perbet
RJUB	6	Justus Randolph
RJV	20	Javier Ruiz Fernandez
SDOH	31	Solar Dynamics Obs - HMI

Continued

Table 3: 202208 Number of observations by observer.

Observer Code	Number of Observations	Observer Name
SNE	5	Neil Simmons
SRIE	24	Rick St. Hilaire
TDE	21	David Teske
TNIA	16	Nick Tonkin
TPJB	2	Patrick Thibault
TST	20	Steven Toothman
URBP	30	Piotr Urbanski
VIDD	17	Dan Vidican
WGI	4	Guido Wollenhaupt
WND	20	Denis Wallian
WWM	29	William M. Wilson
Totals	1359	72

3.4 Generalized Linear Model of Sunspot Numbers

Dr. Jamie Riggs, Northwestern University and Thomas Jefferson University, maintains a relative sunspot number (R_a) model containing the sunspot numbers after the submitted data are scrubbed and modeled by a Generalized Linear Mixed Model (GLMM), which is a different model method from the Shapley method of calculating R_a in Section 3 above. The GLMM is a statistical model that accounts for variation due to random effects and fixed effects. For the GLMM 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. For more details: *A Generalized Linear Mixed Model for Enumerated Sunspots* (see 'GLMM06' in the sunspot counts research page at http://www.spesi.org/?page_id=65).

Figure 8 shows the monthly GLMM R_a numbers for a rolling eleven-year (132-month) window beginning within the 24th solar cycle and ending with last month's sunspot numbers. The solid cyan curve that connects the red *X's* is the GLMM model R_a estimates of excellent seeing conditions, which in part explains why these R_a estimates often are higher than the Shapley R_a values. The dotted black curves on either side of the cyan curve depict a 99% confidence band about the GLMM estimates. The green dotted curve connecting the green triangles is the Shapley method R_a numbers. The dashed blue curve connecting the blue *O's* is the SILSO values for the monthly sunspot numbers. 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 dots below and above the whiskers traditionally are considered outliers, but with GLMM modeling, they are observations that are accounted for by the GLMM model.

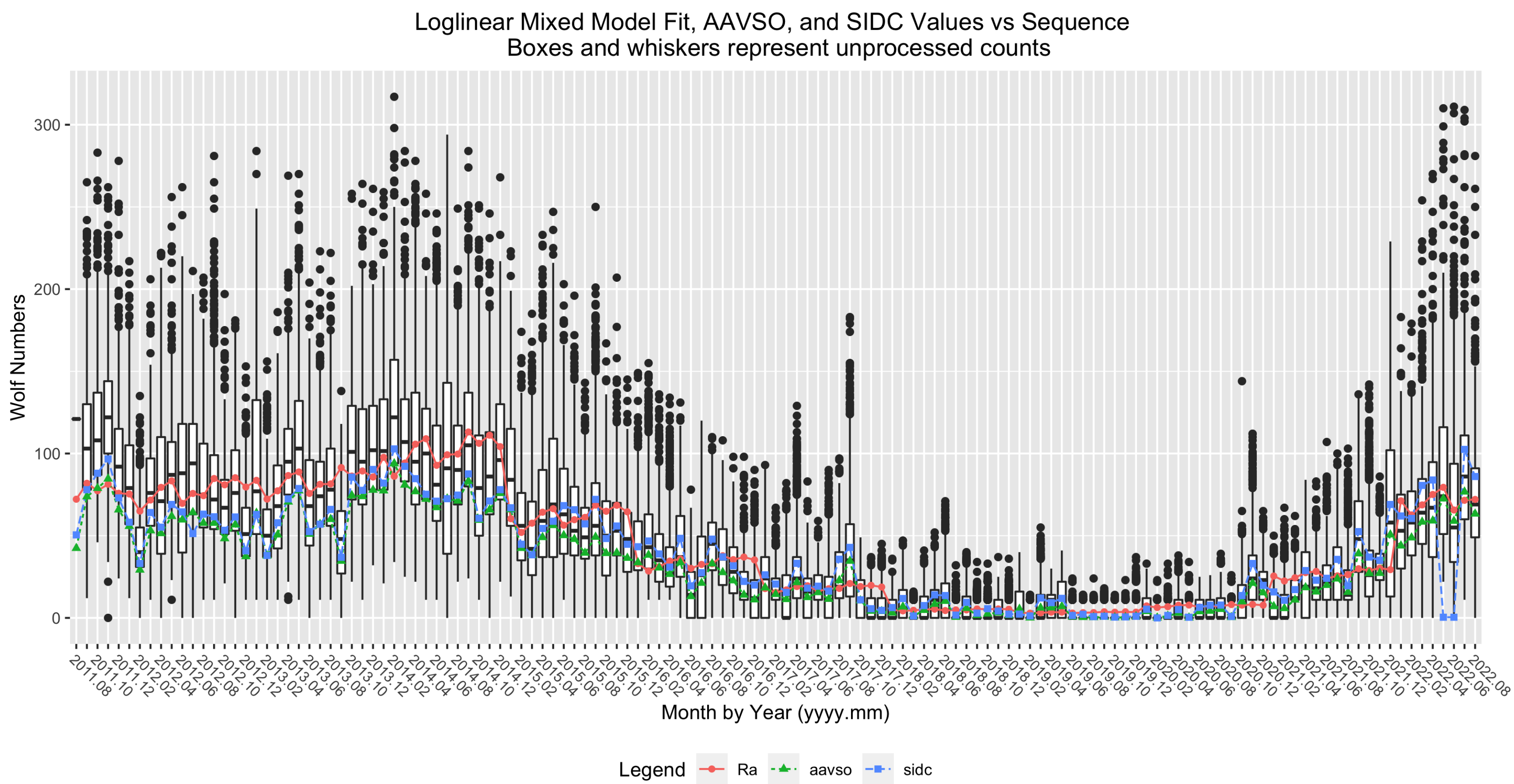
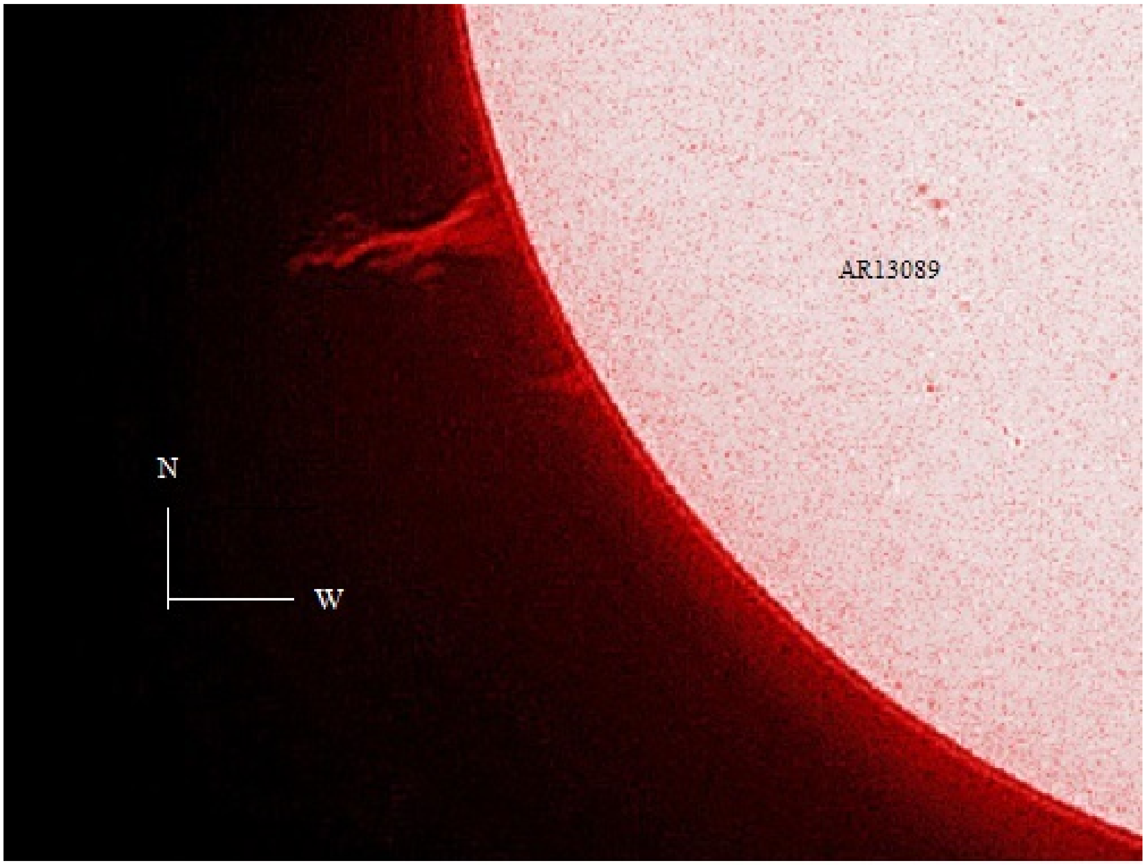


Figure 8: GLMM fitted data for R_a. AAVSO data: https://www.aavso.org/category/tags/solar-bulletin. SIDC data: WDC-SILSO, Royal Observatory of Belgium, Brussels

4 Endnotes

- Sunspot Reports: Kim Hay solar@aavso.org
- SID Solar Flare Reports: Rodney Howe rhowe137@gmail.com



Monty Leventhal OAM. Digital filtergram. Date: 27th August 2022. Time: 23.45 hrs UT.
Telescope: Meade 10". Filter DayStar H-alpha Combo Quark. Camera: Canon EOS 600D.
Exposure 1/8 sec 400 ISO. Supported by the Donovan Astronomical Strust. Testar
Australia Pty. Ltd & the Ridley Grant of the BAA.
Prominence height, Approximately 140K km.

Figure 9: Monty Leventhal (OAM) takes an image of AR3089 early on before it got huge a week later. Perhaps it was the 104,000 km prominence that caused all those flares and VLF SID Events for the 27th of August? (See Figure 3).

Software

The following are the R and R packages used in comparing Sunspot group counts.
R [R Core Team, 2021]: A language and environment for statistical computing.
R package modeltime.gluonts Dancho [2022]: modeltime.gluonts: GluonTS Deep Learning_. R package version 0.3.1, https://github.com/business-science/modeltime.gluonts
R package Tidymodels Kuhn and Wickham [2020]: Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. https://www.tidymodels.org **R package Tidyverse** Wickham et al. [2019]: A collection of R functions.
R package timetk Dancho and Vaughan [2022]: C Tool Kit for Working with Time Series in R_. R package version 2.8.1, https://CRAN.R-project.org/package=timetk

References

AAVSO. AAVSO solar observing section, 2022. URL https://www.aavso.org/solar.

Matt Dancho. *modeltime.gluonts: 'GluonTS' Deep Learning*, 2022. URL https://github.com/business-science/modeltime.gluonts. R package version 0.3.1.

Matt Dancho and Davis Vaughan. *timetk: A Tool Kit for Working with Time Series in R*, 2022. URL https://CRAN.R-project.org/package=timetk. R package version 2.8.1.

Max Kuhn and Hadley Wickham. *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.*, 2020. URL https://www.tidymodels.org.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL https://www.R-project.org/.

SILSO, World Data Center - Sunspot Number and Long-term Solar Observations, 2022. Sunspot number catalogue, 1850-2022 [data set]. Royal Observatory of Belgium. URL https://www.sidc.be/silso/datafiles.

U.S. Dept. of Commerce, NOAA, Space Weather Prediction Center, 2022. URL ftp://ftp.swpc. noaa.gov/pub/indices/events/.

Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4 (43):1686, 2019. doi: 10.21105/joss.01686.