The Astrophysical Journal, 928:97 (5pp), 2022 March 20 https://doi.org/10.3847/1538-4357/ac5963


# Investigation of the Long-term Variation of Solar Ca II K Intensity. II. Reconstruction of Solar UV Irradiance

Jun Kakuwa[1] and Satoru Ueno[2]
[1] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan; kakuwa.jun.a98@kyoto-u.jp
[2] Hida Observatory, Kyoto University, Kurabashira, Kamitakara, Takayama, Gifu 506-1314, Japan; ueno@kwasan.kyoto-u.ac.jp
*Received 2021 December 24; revised 2022 February 20; accepted 2022 February 27; published 2022 March 29*

## Abstract

Reconstruction of long-term solar UV variations during the entire 20th century is reported. The sunspot number has been used for this purpose so far. By using the full-disk Ca K intensity as an additional solar UV proxy, the range of allowed values for the reconstructed UV irradiance becomes more restricted. We use long-term archival data of the photographic Ca K plates digitized at the Kodaikanal Solar Observatory. The photographic calibration method developed in our previous paper (Paper I) is applied. Various long-term proxy data of solar activity have been used to estimate past UV irradiance. In light of this context, some issues using the historical Ca K data are commented on.

*Unified Astronomy Thesaurus concepts:* Astronomy data analysis (1858); Photoheliographs (1230); Solar chromosphere (1479); Solar cycle (1487); Solar ultraviolet emission (1533); Solar-terrestrial interactions (1473)

## 1. Introduction

The solar Ca II K line (393.37 nm) is a broad absorption line located at the short-wavelength end of the optical range. The Ca K line is known as one of the proxies of solar magnetic activity; it is bright in active regions, including plages and enhanced network regions, and in quiet-Sun network lanes in the chromosphere. Solar full-disk Ca K images have been archived around the world since the end of the 19th century in the form of, for example, photographic plates and films (Chatzistergos et al. 2020). Monitoring observations of the full-disk Ca K line have been continued with various spectral widths. Data sets with smaller spectral widths have information on the upper chromosphere. Historical photographic observations with a spectroheliograph tend to have a smaller spectral width than recent observations with a filtergraph.

Solar UV radiation is also bright in active regions. It is known, from early-time observations, that full-disk images of Ca K and some lines in the UV region are visually very similar (Tousey 1967). Continuous UV observations by satellites began around 1980 (DeLand et al. 2019).

Correlating the intensity of Ca K and UV during the overlapping period of these observations, we can indirectly measure solar UV irradiance in the past, before the satellite era. Long-term UV irradiance is significant in the context of Sun–Earth or star–planet relations (Lilensten 2007). Solar UV radiation shows comparatively large quantitative and fractional variations with solar cycles than other wavelengths (Fröhlich & Lean 2004; Woods & DeLand 2021). UV photons with wavelengths $\lesssim$300 nm incident on Earth's atmosphere cause ionization, dissociation, and heating from the ionosphere to the troposphere, depending on the wavelength (Watanabe 1958; Andrews 2000). Because both the interior and exterior regions of the atmosphere affect the atmospheric environment, detailed information on variations of both sides is necessary for the investigation and prediction of various changes in the atmosphere (Jin et al. 2011; Kusano et al. 2021). For example, it is implied that the cooling effect by greenhouse gases on the upper atmosphere may leave traces in the solar-quiet variations of the geomagnetic field (Sq). However, there has been no consensus even on whether Sq was increasing or decreasing during the 20th century after subtracting the solar UV contribution to Sq (Elias et al. 2010; Shinbori et al. 2014; Yamazaki & Maute 2017). One of our tasks is to improve the accuracy of the estimate of UV irradiance in the last century by analyzing long-term data on solar activity.

In this paper, we present the result of reconstructing long-term solar UV irradiance during the 20th century by combining, for the first time, long-term data on the sunspot number, 10.7 cm radio flux (F10.7), and the full-disk Ca K images. Combining different kinds of UV proxies makes it possible to restrict the allowed range for the reconstructed UV irradiance.

Because the digitized photographic plates of Ca K record not the intensity but the photographic density, which are related nonlinearly, it is desirable to convert the density to intensity before applying the photographic images to the UV reconstruction; this is an independent task. We perform the density-to-intensity calibration using a previously proposed method (Kakuwa & Ueno 2021, hereafter Paper I).

The result of our UV reconstruction shows that when the long-term Ca K data are taken into consideration, the reconstructed UV irradiance becomes systematically higher than the case with only the sunspot number at solar minima in the early 20th century. Although using multiple proxies is effective, we emphasize that the procedure of both photographic calibration of the Ca K plates and the construction of the long-term composite UV data for comparison with UV proxies are developing and need further improvement.

## 2. Data

We reconstruct the solar UV irradiance in the last century from long-term data on sunspot number, F10.7 flux, and Ca K intensity by correlating these proxies of solar activity with observed UV data during the overlapping period of observations.

In order to obtain the time series of the Ca K intensity, we analyze long-term archival data of the full-disk Ca K images of the Kodaikanal Solar Observatory (KSO) (Priyal et al. 2014). About 46,000 images of digitized plates from 1904 to 2007, with a spectral width of 0.05 nm, are publicly available online.[3] The raw data are the transparency of the plates. By simply converting them to positive images, we obtain images of the photographic density (Figure 1 in Paper I). Solar cycles are not recognized in the time-series plot of the mean density of the solar disk (Figure 7 in Paper I). This is because the density variations due to solar activity are smaller than daily variations of nonsolar origin (e.g., atmospheric conditions, exposure, and development procedure). Hence, the density-to-intensity calibration is essential.

Let us briefly summarize the calibration procedure (see Paper I for details). We extract quiet regions from a raw density image (Figure 3 in Paper I). Smoothing the image of the quiet regions, we make a background image. From the raw and background density images, an intensity image is obtained by using the density-to-intensity calibration formula (Figure 4 in Paper I). Because of the nonlinear response of the photographic density to the exposure, a simple linear conversion is inadequate in principle (Dainty & Shaw 1974). After iterative improvement in extracting quiet regions and making a background image, we finally obtain an intensity image to be used for long-term analysis. The calibration formula makes the mode of the intensity and the intensity fluctuation of the inner region of the chromospheric network constant with solar cycles through the finally obtained data set (Figure 8 in Paper I). From the data set after the calibration, we can compute the time series of the full-disk Ca K intensity.

There is nonuniformity in the data set of KSO. Priyal et al. (2019) discussed the possibility of wavelength shifts. We have shown in Figure 7 in Paper I that the full-disk mean density begins decreasing at around 1950. Figure 9 in Paper I also shows artificial variations in the raw data set. Statistical variations become larger as well after around 1976.

It is possible to compensate for such artificial variations by uniformly applying the automated calibration procedure described above to the whole data set. But the automated procedure is not always sufficient for the calibration of each image, depending on the quality of the raw data. It is desirable to check the quality of each calibrated image before computing the long-term variation of the Ca K intensity. We visually inspected the calibrated data set and then selected about 10,000 very high-quality images where the plages and chromospheric network are sharply resolved. Although the number of selected images is acceptable, we found that most of them are distributed before 1940. It is therefore impractical to use only the highest-quality images for the long-term UV variation study. Thus, we just remove about 6000 lowest-quality images (e.g., those lacking part of the disk).[4] There remains room for improvement in treating middle-quality data by optimizing the calibration procedure for each image, though the improvement needs a large amount of manpower or time cost and is beyond the scope of this paper. The essential point of the present paper is to demonstrate, for the first time, that historical data on the full-disk Ca K line, which have been archived around the world, have the potential to reconstruct more reliable solar UV irradiance in the last century.

Regarding the UV data to be compared with the sunspot number, F10.7, and the Ca K intensity, we use the SSI3 composite[5] available on the LASP Interactive Solar Irradiance Data Center (Woods & DeLand 2021). The SSI3 composite is the time series of the solar UV irradiance from 1978 to 2019 with a wavelength bin of 1 nm. In general, a long-term UV data set is composed of data from multiple satellites, and therefore, their calibration is another significant challenge in long-term variation studies (Deland & Cebula 2008; Yeo et al. 2014; Chamberlin et al. 2020; Woods & DeLand 2021). Because the wavelength range covered by UV satellites has been gradually widened, long-term UV data sets usually consist of not only actually observed data but also other proxy data (e.g., Mg II and Ly$\alpha$). We use all the data of the SSI3 composite without any selection in order to make a longer-time comparison with our UV proxies in a wider wavelength range. Future improvement of the calibration of the UV data will make the UV reconstruction more reliable.

Data on the sunspot number[6] and F10.7[7] can also be downloaded from LISIRD. The data from 1818 and 1947 are available for the former and the latter, respectively.

## 3. Results

Figure 1 shows the scatter plot between the three UV proxies and the UV irradiance after 1978, where the 12 month moving median of each data set is plotted. UV irradiance is integrated in three wavelength ranges: 0.5–133.5, 134.5–199.5, and 200.5–299.5 nm. The figure displays only the second case; the other two are omitted because there is no noteworthy difference (see Table 1 instead). For the moving median of the UV data in each panel, only the UV data accompanied by the proxy data on the same day are used. The number of such UV data is about 2600 for the Ca K intensity, while for F10.7 and the sunspot number, it is about 15,000, which is much larger than Ca K.

Each of the three proxies of solar activity is complementary for reconstructing the UV irradiance. It has been known observationally that the F10.7 flux at solar minima is extremely stable if compared among different solar cycles (Shimojo et al. 2017). The nonlinearity between the F10.7 and UV flux is also known[8] (Balan et al. 1994). The sunspot number tends to become zero by definition at solar minima, as can be seen in the bottom panel of Figure 1. These characteristics of F10.7 and the sunspot number can lower the sensitivity of these proxies to the UV variations at solar minima and/or maxima in particular. On the other hand, the decades-long relation between Ca K intensity and spectroscopic UV irradiance is not well-known quantitatively in detail (but see earlier works by Lean et al. 1982 and Worden et al. 2001).

[3] https://kso.iiap.res.in/new/

[4] See Paper I for a sample of the highest-quality images. We omit showing samples of the lowest-quality images as there is a variety of cases.

[5] https://lasp.colorado.edu/lisird/data/lasp_gsfc_composite_ssi/

[6] https://lasp.colorado.edu/lisird/data/international_sunspot_number/

[7] https://lasp.colorado.edu/lisird/data/penticton_radio_flux/

[8] The nonlinearity is not seen in Figure 1, which shows the data smoothed in time and integrated in wavelength. This could be due to the 12 month smoothing that averages out the solar rotation effects and thus highlights only the 11 yr solar cycle effects. According to Woods et al. (2000), both the dominant plage components and the active network components are needed for modeling UV irradiance. They show that these two solar features manifest themselves differently for the 27 day solar rotation variability and the 11 yr solar cycle.

**Table 1**
Parameters of the Regression Line between Each Proxy and Integrated UV Irradiance

| UV Wavelength Range (nm) | Ca K $a$ | $b$ | $\sigma$ | F10.7 $a$ | $b$ | $\sigma$ | Sunspot Number $a$ | $b$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.5–133.5 | 42.6 | −30.4 | 1.15 | 0.048 | 7.4 | 0.41 | 0.031 | 10.45 | 0.48 |
| 134.5–199.5 | 37.75 | 60.27 | 1.14 | 0.045 | 93.4 | 0.46 | 0.0294 | 96.26 | 0.54 |
| 200.5–299.5 | 818 | 14295 | 34.8 | 1.015 | 15005 | 16.4 | 0.626 | 15072 | 18.3 |

**Note.** From a value of each proxy ($x$), the UV irradiance ($y$) is computed as $y = ax + b$. Units of $x$ and $y$ are the same as in Figure 1. The statistical error of this conversion is $\sigma$ (see Section 3).

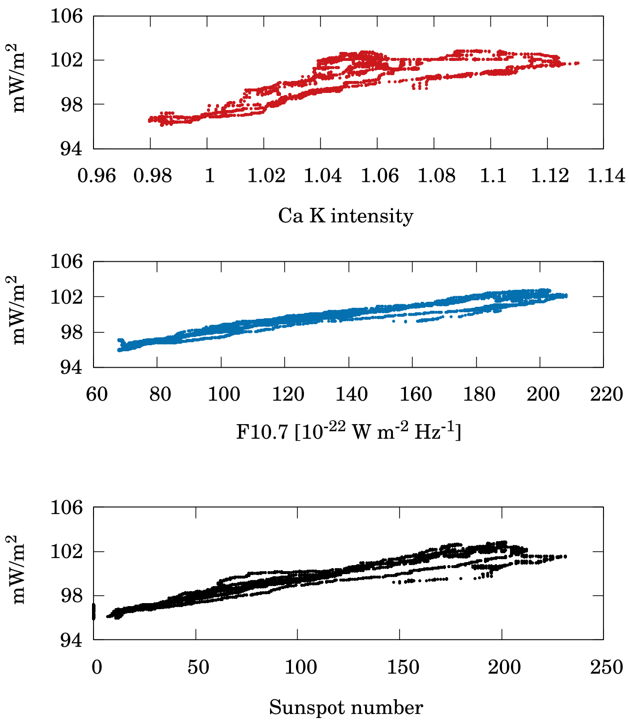


**Figure 1.** Scatter plot between each UV proxy and the UV irradiance integrated from 134.5 nm to 199.5 nm. The 12 month moving median of each data set after 1978 is plotted. The Ca K intensity is normalized by the mean intensity of quiet regions near the disk center.

Our present analysis of the KSO data set shows a comparatively large scatter (Figure 1; see also Table 1). This is probably due to variations in the quality of raw data and/or effectiveness of the photographic calibration procedure for each plate. For example, the variations of the photographic density fluctuation in solar-quiet regions due to seeing conditions affect the result. The uncertainty of the relation between the density and the intensity for active regions also has an influence (see Paper I for discussion).

The top panel of Figure 1 may indicate two families. We cannot make sure whether it does appear only in the top panel. But the difference in the separation degree from the other two panels could be due to either influence of middle-quality data or inadequate photographic calibration for a subset of data.

We reconstruct the long-term solar UV irradiance not from a single proxy but from multiple proxies. Table 1 shows the parameters of regression lines between the 1 month moving median of proxy data and the UV data. (Note that rougher (i.e., 12 month) smoothing is applied to Figure 1, which helps us see the trend of the relation over a timescale longer than a solar cycle.) The statistical errors of the UV reconstruction with the regression lines are quantified by the mean absolute deviation between the UV data and the lines, $\sigma$. About 60% of the UV data are located within $\pm\sigma$ from the line for all the cases. For the Ca K intensity, $\sigma$ is comparatively larger than the other two proxies. The value of $\sigma$ will become a benchmark for future improvement of processing historical Ca K data.

We convert the three proxies to UV irradiance using the regression lines. The parameters are assumed to be constant in time. The uncertainty of the UV reconstruction by each proxy is taken into consideration through $\sigma$. In Figure 2 we show the reconstructed solar UV irradiance before 1978, where there is no data of the SSI3 composite. The black line displays the result from only the sunspot number. The line width indicates the uncertainty $\sigma$.

The true value of the UV irradiance is probabilistically expected to lie within the reconstruction uncertainty $\sigma$ of each proxy; this applies regardless of the magnitude of $\sigma$. We show with the green line the result consistent among all the three proxies. This is obtained as the area of overlap among the results of the three independent proxies. This is the main result of the present paper. The line width indicates the range of the reconstructed UV irradiance allowed by the three proxies. The width varies with time, having the average half-width of 0.38, 0.43, and 15.6 for the three UV wavelength ranges from short to long ones, respectively.[9] Only the results for 134.5–199.5 nm are shown in the figure as in Figure 1. The results for the other two wavelength ranges resemble the case shown and are omitted. We will comment on this in Section 4.

The sunspot number has been used as a UV proxy covering the 20th century. The result of our UV reconstruction with the three proxies suggests that the solar UV irradiance is systematically higher than the estimate by only the sunspot number, particularly at solar minima before 1940. The result from only the sunspot number is probably affected by its artificial behavior at solar minima as mentioned before in this section.

F10.7 can be used as a proxy after 1947. F10.7 is taken into consideration in our result in Figure 2 (green line), though it is practically ineffective because of its similar behavior to the sunspot number.

It is not straightforward to compare the reconstructed UV irradiance with other long-term UV models. Different UV models generally refer to different observational UV data sets to reconstruct long-term UV irradiance. In addition, in the case where UV data sets provided by several instruments are used in a UV model, how they are synthesized is different for different models. For example, SSI3 and NRLSSI2[10] (Coddington et al. 2016) show different flux levels and long-term variation

[9] The flux level of UV irradiance is higher for the longer-wavelength ranges. The absolute value of the uncertainty gets larger for longer-wavelength ranges, while the relative uncertainty gets smaller.

[10] https://lasp.colorado.edu/lisird/data/nrl2_ssi_P1D/

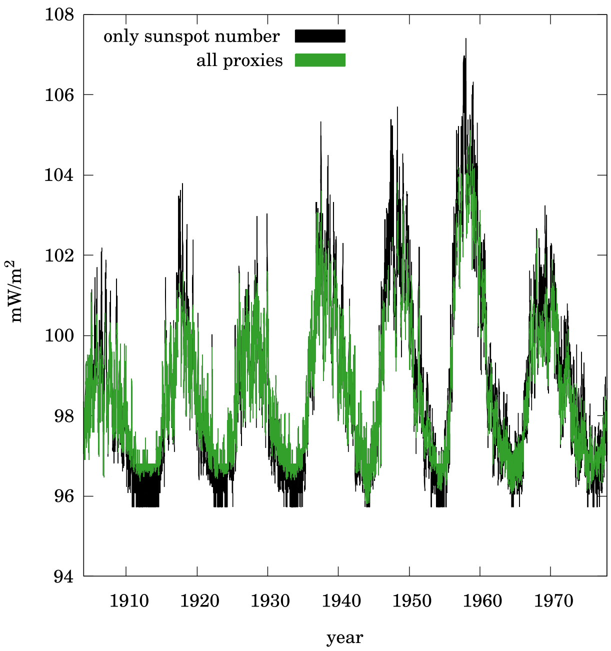


**Figure 2.** Reconstructed solar UV irradiance variations from 1904 to 1978. The results for the wavelength range of 134.5–199.5 nm are shown. The line width indicates the uncertainty. The green line and the black line are the results obtained from the three proxies (i.e., the sunspot number, F10.7, and Ca K) and only the sunspot number, respectively.

amplitudes. Hence, the calibration of long-term UV data is clearly an important subject, though it is beyond the scope of this paper; comparison between SSI3 and other models is not the purpose. If we use UV models other than SSI3 for the UV reconstruction, the resulting values of the UV irradiance will more or less change. But the significance of developing a new method of extracting the Ca K intensity information from historical photographic data and applying it to the UV reconstruction is unchanged.

## 4. Discussion and Summary

In this paper, we have reconstructed solar UV irradiance variations during the 20th century using the long-term data from full-disk Ca K images together with two other conventional proxies: the sunspot number and F10.7. The reconstructed UV irradiance shows systematic differences between the cases with and without the Ca K data, particularly in the early 20th century (Figure 2). This is not surprising because of the characteristics of the sunspot number and F10.7 (Section 3). However, how the results with and without the Ca K information differ needs further investigation as described below.

As mentioned in Section 2, the data quality of the digitized Ca K images of KSO probably decreased toward the end of the 20th century. We have correlated the Ca K data after 1978 with the UV irradiance data. We cannot deny at all the possibility that such changes in the quality affect the UV reconstruction, though it must be more or less absorbed by the photographic calibration procedure. Unfortunately, if we select only the highest-quality data, we will not have enough data for the long-term variation study (Section 2). It is unclear to what extent middle-quality data affect the results. Although the uncertainty of the UV reconstruction with each proxy is quantitatively taken into consideration by $\sigma$, there remains the possibility that future development of more sophisticated photographic calibration techniques changes the results.

Actually, there have been some indications that analyses of historical Ca K plates in some previous studies were affected by the quality changes. For example, there are a lot of previous studies analyzing the long-term variations of plage areas from Ca K plates. Different from the present paper challenging photographic calibration and analyzing full-disk Ca K intensity variations, analysis of the plage area finally requires binary images in which active regions and quiet regions are given the values of 1 and 0, respectively. In such analysis, the influence of quality changes should be minimized, and hence it can utilize the maximum number of data for long-term studies. Even in this case, however, different results have been obtained from the same data set of KSO. The time series of the plage area reaches maximum value at solar cycle 19 in most cases, but the difference in the maximum value between cycles 19 and 21 has a variety of cases, and cycle 21 shows the maximum value in an extreme case (Chatterjee et al. 2016; Priyal et al. 2017; Chatzistergos et al. 2019). This inconsistency is probably due to both the nonuniformity of the KSO data set (Priyal et al. 2019) and the difference in automated calibration procedures, showing the importance and difficulty of data selection and calibration for long-term analysis.

One solution is to increase high-quality samples, especially after around 1940–1950 (Section 2). Photographic Ca K images are archived at the National Astronomical Observatory of Japan as well. A remarkable point is that the spectral width is the same as in KSO so it will be reasonable to combine the two data sets. Another issue is to optimize the photographic calibration procedure for each plate.

For the validity test of the Ca K intensity analysis, it is necessary to accumulate modern Ca K intensity data with the same spectral widths as historical photographic data. Comparison with traditional plage area analyses does not work as the validity test. The Ca K intensity integrated over the full solar disk is computed from the information on the size and brightness of active regions, while the plage area computation needs only the former. In addition, different Ca K data sets generally have different spectral widths, which mean different observed heights in the solar atmosphere and thus result in different relative intensity values between the active and quiet regions.

As for the UV reconstruction, the comparison of the intensity of Ca K and UV is physically more natural than the comparison between the plage area and the UV intensity. Although the intensity analysis can suffer from lower-quality data, analyses of the plage area in previous studies also show inconsistencies, as mentioned before in this section (see also Section 4 in Paper I). From a general viewpoint, what is important for UV reconstruction is to finally deduce a UV irradiance consistent across multiple proxies within the statistical reconstruction uncertainty of each proxy (Figure 2) rather than which proxy we select.

In this study, we have reproduced solar UV variations, dividing UV into three wavelength ranges. But we cannot evaluate which is the most adapted waveband for our reconstruction procedure because of the limitation in the accuracy of currently published UV composite data. For example, in the SSI3 data set, only the 250–340 nm waveband actually uses UV data observed by Nimbus-7/SBUV from the beginning of 1978. Satellite observations at a shorter wavelength than 115 nm have only existed since the 2000s.

Thus, UV irradiance values of the composite data are interpolated using proxies, such as Mg II and Ly$\alpha$, widely in both wavelength and time period when actual observed data do not exist. Another UV spectroscopic data set, FISM2 (Chamberlin et al. 2020), also has the same problem of using proxy data instead of satellite observations. For a more robust reconstruction, it is necessary to wait for sufficient accumulation of modern Ca K spectroscopic data with the same spectral widths as historical photographic observations after the 2000s with simultaneous spectroscopic UV observations.

## ORCID iDs

Jun Kakuwa https://orcid.org/0000-0001-8428-5071
Satoru Ueno https://orcid.org/0000-0002-2165-7998

## References

Andrews, D. G. 2000, An Introduction to Atmospheric Physics (Cambridge: Cambridge Univ. Press)
Balan, N., Bailey, G. J., Jenkins, B., Rao, P. B., & Moffett, R. J. 1994, JGR, 99, 2243
Chamberlin, P. C., Eparvier, F. G., Knoer, V., et al. 2020, SpWea, 18, e02588
Chatterjee, S., Banerjee, D., & Ravindra, B. 2016, ApJ, 827, 87
Chatzistergos, T., Ermolli, I., Krivova, N. A., et al. 2020, A&A, 639, A88
Chatzistergos, T., Ermolli, I., Krivova, N. A., & Solanki, S. K. 2019, A&A, 625, A69
Coddington, O., Lean, J. L., Pilewskie, P., Snow, M., & Lindholm, D. 2016, BAMS, 97, 1265
Dainty, J. C., & Shaw, R. 1974, Image Science. Principles, Analysis and Evaluation of Photographic-type Imaging Processes (London: Academic Press)
Deland, M. T., & Cebula, R. P. 2008, JGRA, 113, A11103
DeLand, M. T., Floyd, L. E., Marchenko, S., & Tiruchirapalli, R. 2019, E&SS, 6, 1284
Elias, A. G., de Artigas, M. Z., & de Haro Barbas, B. F. 2010, JGRA, 115, A08316
Fröhlich, C., & Lean, J. 2004, A&ARv, 12, 273
Jin, H., Miyoshi, Y., Fujiwara, H., et al. 2011, JGRA, 116, A01316
Kakuwa, J., & Ueno, S. 2021, ApJS, 254, 44, (Paper I)
Kusano, K., Ichimoto, K., Ishii, M., et al. 2021, EP&S, 73, 159
Lean, J. L., White, O. R., Livingston, W. C., et al. 1982, JGR, 87, 10307
Lilensten, J. 2007, Space Weather : Research Towards Applications in Europe (Dordrecht: Springer)
Priyal, M., Singh, J., Belur, R., & Rathina, S. K. 2017, SoPh, 292, 85
Priyal, M., Singh, J., Ravindra, B., Priya, T. G., & Amareswari, K. 2014, SoPh, 289, 137
Priyal, M., Singh, J., Ravindra, B., & Shekar, B. C. 2019, SoPh, 294, 131
Shimojo, M., Iwai, K., Asai, A., et al. 2017, ApJ, 848, 62
Shinbori, A., Koyama, Y., Nose, M., et al. 2014, EP&S, 66, 155
Tousey, R. 1967, ApJ, 149, 239
Watanabe, K. 1958, AdGeo, 5, 154
Woods, T. N., & DeLand, M. T. 2021, E&SS, 8, e01740
Woods, T. N., Tobiska, W. K., Rottman, G. J., & Worden, J. R. 2000, JGR, 105, 27195
Worden, J. R., Woods, T. N., & Bowman, K. W. 2001, ApJ, 560, 1020
Yamazaki, Y., & Maute, A. 2017, SSRv, 206, 299
Yeo, K. L., Krivova, N. A., & Solanki, S. K. 2014, SSRv, 186, 137