Journal of the Korean Astronomical Society
47: 255 ∼ 258, 2014 December

http://dx.doi.org/10.5303/JKAS.2014.47.6.255
pISSN: 1225-4614 · eISSN: 2288-890X
http://jkas.kas.org

Statistical Study on Personal Reduction Coefficients of Sunspot Numbers Since 1981

Il-Hyun Cho[1,2], Su-Chan Bong[1,2], Kyung-Suk Cho[1,2], Jaejin Lee[1,2], Rok-Soon Kim[1], Young-Deuk Park[1], and Yeon-Han Kim[1,2,3]

[1]Korea Astronomy and Space Science Institute, 776 Daedeok-daero, Yuseong-gu, Daejeon 305-348, Korea
yhkim@kasi.re.kr
[2]University of Science and Technology, Gajung-ro, Yuseong-gu, Daejeon 305-350, Korea
[3]Big Bear Solar Observatory, New Jersey Institute of Technology, Big Bear City, CA 92314, USA

Received August 18, 2014; accepted October 20, 2014

Abstract: Using sunspot number data from 270 historical stations for the period 1981-2013, we investigate their personal reduction coefficients (k) statistically. Chang & Oh (2012) perform a simulation showing that the k varies with the solar cycle. We try to verify their results using observational data. For this, a weighted mean and weighted standard deviation of monthly sunspot number are used to estimate the error from observed data. We find that the observed error (noise) is much smaller than that used in the simulation. Thus no distinct k-variation with the solar cycle is observed contrary to the simulation. In addition, the probability distribution of k is determined to be non-Gaussian with a fat-tail on the right side. This result implies that the relative sunspot number after 1981 might be overestimated since the mean value of k is less than that of the Gaussian distribution.

Key words: sunspot: personal reduction coefficient — method: data analysis

1. Introduction

Wolf initiated the notion of relative sunspot number ($10g + f$) by separately defining the group number (g) and individual number (f) to better describe solar activity with great importance on a spot-building action compared to a variation of individual spot (Izenman 1983; Svalgaard 2012). The sunspot number of each observing station indeed differs from each other. Therefore, a standard sunspot count might require a correction factor k that scales the number to a primary number which is so called the Wolf number or Zürich number defined by $k(10g+f)$ (Hoyt & Shatten 1998). Since 1981, the Solar Influence Data Analysis Center (SIDC) has been producing the International Sunspot Number (ISN) continuously over the 35 years from 270 stations (Clette et al. 2014). Since then, the correction factor k becomes a personal reduction coefficient of an arbitrary station which scales its number to the reference number, ISN (Kim et al. 2003; Clette et al. 2007).

The personal reduction coefficient (k) could have an intrinsic dependency on the phase of solar cycle because of its definition as the ratio of the reference number to the number from an arbitrary station. One can expect that the deviation of k from 1.0 is larger during a phase of minimum solar activity than the deviation in maximum phase. For example, a difference in the raw sunspot numbers between an arbitrary and reference station with relatively small values of the numerator and denominator in a minimum phase leads to large deviations from 1.0, while the ratio with relatively large values of the numerator and denominator in a maximum phase leads to small deviations from 1.0. Therefore, k is likely to be coupled with the solar cycle.

Meanwhile, Chang & Oh (2012) investigated whether the variation in the correction factor k (or the personal reduction coefficient) is coupled with the solar cycle using artificial data-sets. They considered two kinds of error distributions, uniform and exponential one, and convincingly showed that k increases as sunspot numbers increase and the variation is independent on the statistical distribution of error. They also showed that the amplitude of k-variation in the solar cycle might be positively correlated with the signal to noise ratio (S/N) of the data.

We aim to study the statistics of the personal reduction coefficient by using all contributing stations from 1981. We also try to examine the distribution of k which could be related to the distribution of sunspot errors among observers. This might be helpful to understand the cycle dependency of the personal reduction coefficient as reported by Chang & Oh (2012).

2. Method and Result

The definition of the personal reduction coefficient is $k = R_L/R_{A,raw}$, where the R_L represents the reference value observed from the Locarno station, scaled by a factor of 0.6 which is a scaling factor between the current R_L and previous Zürich number (Clette et al. 2007). The $R_{A,raw}$ represents a raw sunspot number observed by an arbitrary station. If daily k deviate by more than $2\sigma_k$ from the monthly average of k, they are discarded from the data-set and *monthly* k are re-calculated until the average does not change (Clette et al. 2007). If the Locarno station did not

Corresponding author: Y.-H. Kim

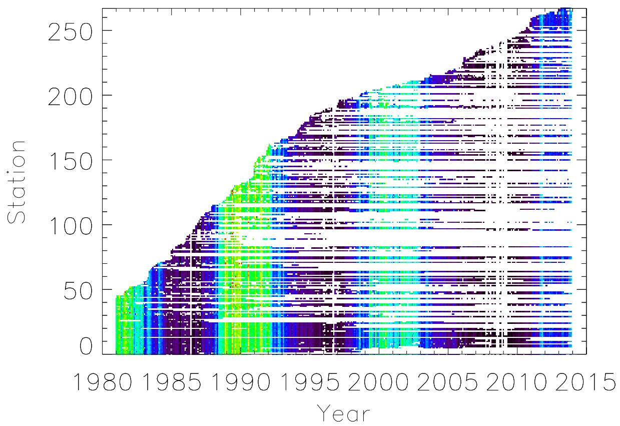

Figure 1. Coverage of sunspot observations for all contributing stations. Numbers in the y-axis represent arbitrary stations. The color indicates the sunspot numbers. Black-violet parts approximately represent solar minimum periods. Each point is monthly sampled. The white blank corresponds an epoch of an insufficient number (lower than 7-days for a month) of observations or no observations.

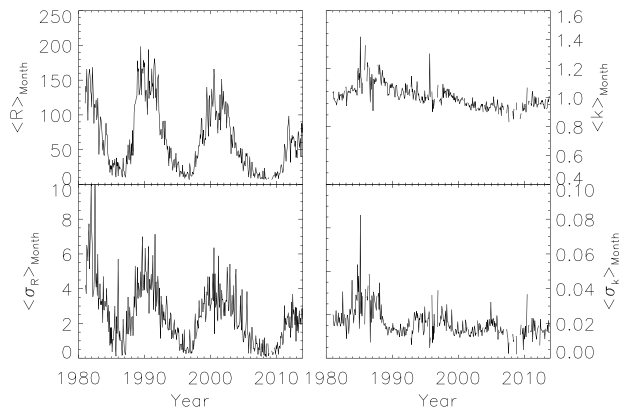

Figure 2. The weighted average of monthly relative sunspot numbers (upper left) observed by 270 stations since 1981, personal reduction coefficients (upper right), and corresponding weighted standard deviations (lower left and right).

see the Sun due to a bad weather condition or an instrumental problem, the R_L of that day is replaced by the network average of relative sunspot number (R) which is scaled by their own historical k. After the *monthly* k is calculated, daily relative R_A are determined by multiplying with their *monthly* k. Figure 1 represents coverage of sunspot observations during 1981-2013 for all stations. One can see that the monthly number of stations as a function of time is not uniform which means that a sample at a certain month might be measured with better or worse precision than others (Bevington & Robinson, 2002; Usoskin et al. 2003). Therefore, we apply a weighted mean and standard deviation to better estimate the relative sunspot number and its uncertainty. The weighted average at each month (μ) is given by $\mu = \Sigma(x_i/\sigma_i^2)/\Sigma(1/\sigma_i^2)$, where σ_i is uncertainty of an arbitrary station at each month. The weighted standard deviation (σ_μ) at each month is given by $\sigma_\mu^2 = 1/\Sigma(1/\sigma_i^2)$.

Figure 2 shows monthly weighted average of relative sunspot number and its personal reduction coefficient together with their weighted standard deviations. The reason we use the weighted moments is that the number of station is not uniformly distributed as mentioned above. This non-uniformity might be an intrinsic property of the sunspot count though one can only assume its error property by comprehensive model (Chang 2008, 2009). The weighted standard deviation of the observed sunspot number exhibits the coupling with the solar cycle as seen in the lower left panel of Figure 2 that was similar to those assumed in previous studies (Chang 2008, 2009; Chang & Oh 2012). It is seen that the personal reduction coefficient shows locally higher values at the cycle minima. A slow decrease is also observed over three decades. Clette et al. (2014) suspect that it might be due to the eyesight degradation and aging effect of the lead observer at the reference sta-

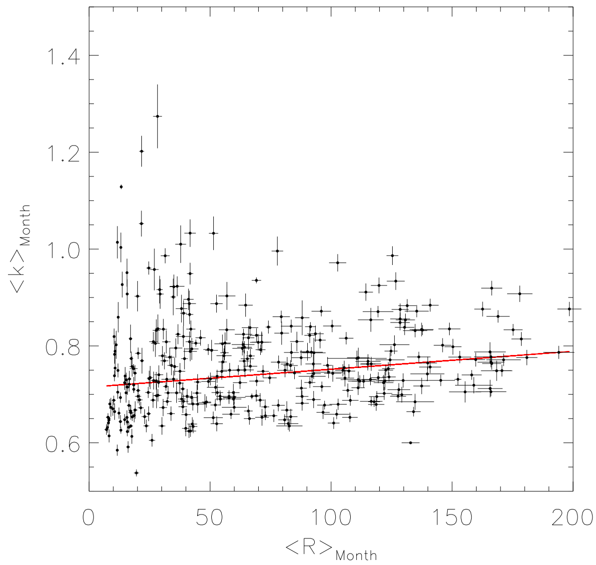

Figure 3. Scatter plot of the personal reduction coefficient as a function of the sunspot number. Red-straight line represents least square fit which shows a weak-positive trend as the sunspot number increases. Horizontal and vertical bars indicate 1-sigma of the relative sunspot numbers and personal reduction coefficient at each value.

tion. Estimated error or noise from observed data (the inverse of signal to noise ratio; $(S/N)^{-1}$) is approximately 0.0013 which is to be compared to the much larger value of 0.01 considered by Chang & Oh (2012). Our $(S/N)^{-1}$ is directly calculated from $\sigma_{\sigma_r}^2/\sigma_R^2$, where the σ_{σ_r} is the standard deviation of monthly errors from 270 stations and σ_R is the weighted standard deviation for monthly sunspot number.

To examine the k-coupling with the solar cycle more precisely, we plot the personal reduction coefficient as a function of the sunspot number in Figure 3. A red-straight line represents least square fit to the dependency. Although, the personal reduction coefficient is weakly correlated with the sunspot number, this dependency is not likely to be related to the 11-year solar cycle since there is large scatter compared to the dependency variation (the red-straight line) over the whole sample. Note that the amplitude of the 11-year solar cycle shown in the upper panel in Figure 2 is much larger than the change through the solar cycle. The simulated $(S/N)^{-1}$ in the previous study is 0.01 which is much larger than the $(S/N)^{-1}$ estimated in our study (0.0013). Moreover, the former considered intrinsic errors in simulated sunspot number, while the latter used total errors including the intrinsic error in the observed relative sunspot number. Therefore, the intrinsic error in the observed sunspot number is definitely much smaller than 0.01. It seems that the observed k-coupling with the solar cycle might be very small to detect. We suspect that the observed intrinsic error is too small to obtain a visualization of the k-coupling, although the variation of error could be related to the k-coupling itself as seen in the lower left panel of Figure 2 and the assumption in Chang & Oh (2012).

In Figure 4, we plot the distribution of the personal reduction coefficient. The distribution shows a fat-tailed feature in the right side which is likely to be a log-normal or gamma distribution rather than purely Gaussian. This non-Gaussianity is likely to be independent of the length of data-sets in obtaining k implying that it might be an intrinsic property of k rather than a transient due to insufficient data samples. This means that the average and standard deviation used in the 2-sigma criterion of the calculation likely overestimate the average and standard deviation. We estimate the actual average and standard deviation by considering the shape of the distribution. Those averages and standard deviations are represented by red-solid and red-dashed lines, respectively. The average and standard deviation under the assumption of the pure Gaussian is represented by black-solid and black-dashed lines, respectively. It is noted that the former is smaller than the latter by about 11% possibly implying that the relative sunspot number is overestimated by at most 11%.

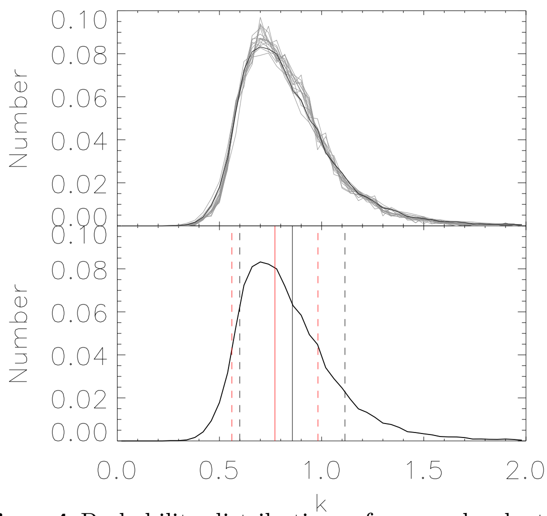

Figure 4. Probability distributions of personal reduction coefficients from 270 stations. Gray curves are obtained by changing the length of data-sets in obtaining k with 10, 20, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, 360, and 720 days. Black curves in upper and lower panel are the probability distribution of monthly personal reduction coefficient. The right side of the distribution is elongated meaning that the distribution is apart from the Gaussian. Black-solid and black-dashed lines indicate a mean and $\pm$1-sigma (0.257) from the mean (0.857) which are calculated under the assumption that the distribution is purely a Gaussian. Red-solid and red-dashed lines indicate a mean (0.772) and $\pm$1-sigma (0.211) from the mean. They are calculated by considering the observed shape of the distribution such that they occupy 50%, 15.87%, and 84.13% area, which correspond a mean and $\pm$1-sigma in the Gaussian distribution.

3. Summary and Discussion

We have examined the variation of the personal reduction coefficient with the solar cycle simulated by Chang & Oh (2012). The variation of k with the solar cycle was not identified in the observed data-sets since the actual error is much smaller than 0.01 used in the simulation. On the other hand, we found that the personal reduction coefficient is weakly correlated with the sunspot number of which the origin is currently unknown.

The distribution of the personal reduction coefficient was first obtained by using data from 270 stations since 1981. The distribution is different from a Gaussian and conserves its non-Gaussianity with the changing length of data-sets in obtaining k, indicating that its statistical moments are likely different from those of Gaussian sample. We estimated the average and standard deviation by considering this non-Gaussian property and suspect that the sunspot number is likely overestimated by at most 11%. In actual case, the calculation process of the relative sunspot number excludes samples located outside of the 2-sigma iteratively while we exclude them only once, a significant portion of the overestimated amount might be reduced. More detailed quantification is beyond the scope of our study.

Acknowledgments

We are grateful to the referee for critical comments which helped significantly to improve our manuscript. This work was supported by the operation of Korean Space Weather Prediction Center (KSWPC) at Korea Astronomy and Space Science Institute (KASI). We are appreciate the WOLF system which integrates individual sunspot numbers from all contributing stations from 1981. KASI station, Solar and Space Weather Research Group, have been registered since 2009 and maintained by KSWPC at KASI.

References

Bevington, P., & Robinson, D. K. 2002, Data Reduction and Error Analysis for the Physical Sciences, 3rd, McGraw-Hill, 59

Chang, H.-Y. 2008, Stochastic Properties in North-South Asymmetry of Sunspot Area, New Astron., 13, 195

Chang, H.-Y. 2009, Periodicity of North-South Asymmetry of Sunspot Area Revisited: Cepstrum Analysis, New Astron., 14, 133

Chang, H.-Y., & Oh, S.-J. 2012, Does Correction Factor Vary with Solar Cycle?, JASS, 29, 97

Clette, F., Berghmans, D., Vanlommel, P., Van der Linden, R. A. M., Koeckelenbergh, A., & Wauters, L. 2007, From the Wolf number to the International Sunspot Index: 25 years of SIDC, ASR, 40, 919

Clette, F., Svalgaard, L., Vaquero, J. M., & Cliver, E. W. 2014, arXiv:1407.3231

Hoyt, D. V., & Schatten, K. H. 1998, Group Sunspot Numbers: a New Solar Activity Reconstruction, Sol. Phys., 181, 491

Izenman, A. J. 1983, J. R. Wolf and H. A. Wolfer: An historical note on the Zurich sunspot relative numbers, J. R. Statist. Soc., 146, 311

Kim, R.-S., Cho, K.-S., Park, Y.-D., Moon, Y.-J., Kim, Y.-H., & Sim, K.-J. 2003, The Relative Sunspot Numbers from 1987 to 2002, PKAS, 18, 25

Svalgaard, L. 2012, How Well Do We Know the Sunspot Number?, Comparative Magnetic Minima: Characterizing quiet times in the Sun and Stars, Proc. IAU, 286, 27

Usoskin, I. G., Mursula, K., & Kovaltsov, G. A. 2003, Reconstruction of Monthly and Yearly Group Sunspot Numbers From Sparse Daily Observations, Sol. Phys., 218, 295