Alternate approach to the analysis of solar photometer data

B. M. Herman, M. A. Box, J. A. Reagan, and C. M. Evans

The standard technique of analyzing solar photometer data to determine atmospheric optical depth and the spectral solar constant is shown to inadvertently weight the data unequally. A new approach is proposed which equally weights all the data. Assuming that the deviations of the data points result from real random variations of optical depth during the period of the measurements, this latter approach is shown to yield more reliable results.

I. Introduction

Solar radiometry is a technique that has become more or less a standard method of monitoring the total atmospheric extinction along a vertical column through the atmosphere. A by-product of the technique is the ability to determine the incident solar flux at zero air mass (i.e., the solar constant) at the wavelength in question. Indeed, it was by this method that the original solar constant determinations[1,2] were made. The method has been widely used in more recent years,[3] and although the instrumentation has become more refined the basic data analysis techniques have remained the same.

Analysis of solar radiometry data is based upon the assumption that the transmission of the direct solar flux follows Beer's law, given by

$$F = F_0 \exp(-m\tau), \tag{1}$$

where F is the directly transmitted solar flux at a particular wavelength λ, F_0 is the solar flux incident at the top of the atmosphere, τ is the vertical optical depth of the entire atmosphere at the wavelength λ, and m is the air-mass correction factor to account for the actual slant path of the sun through the atmosphere. (Throughout this paper, we suppress the wavelength dependence to avoid a plethora of subscripts.) The actual experimental expression is $K \cdot F = K \cdot F_0 \exp(-m\tau$, where K is an instrumental constant. However, in the work that follows, since only relative values of F and F_0 are required, the constant K will be dropped.

In Eq. (1), F is measured, and m is determined from time of measurement; F_0 and τ are the unknowns to be determined. In principle, then, measurements of F at two different values of air mass m (i.e., at different times in the day) should suffice to determine the two unknowns. In practice, however, many measurements are usually made throughout the day, and assuming horizontal homogeneity and no temporal changes of vertical optical depth during the measurements, the resulting data are analyzed for the best value of F_0 and τ in some statistical manner. The most common procedure is to write Eq. (1) in logarithmic form, yielding

$$\ln F = \ln F_0 - \tau m \tag{2}$$

which on a plot of $\ln F$ vs m is a straight line of slope $-\tau$ and intercept $\ln F_0$. The data are then usually analyzed to yield the best-fit straight line in a least-squares sense.

Analyses of the various sorts of errors associated with the U. Arizona solar radiometer indicate that the largest error is, in fact, caused by actual random variations in the total optical depth during the course of the day. These variations can range from <1% of τ on very stable days to 10% or more of τ on other occasions, typically being in the 1–5% range of τ. Other error sources are generally <0.5% of τ. Therefore, we shall here assume that the only error of significance is caused by random fluctuations in τ.

From Eq. (2), it can readily be seen that a fluctuation $m\Delta\tau$ will yield a fluctuation in $\ln F, \Delta \ln F$, given by

$$|\Delta \ln F| = m\Delta\tau. \tag{3}$$

Thus, it can be seen that random deviations in τ will cause deviations of $\ln F$ from a straight line, and that these deviations increase with increasing air mass. Therefore, a simple (unweighted) least-squares fit of the

All authors are with University of Arizona, Tucson, Arizona 85721: J. A. Reagan is in the Department of Electrical Engineering, and the other authors are in the Institute of Atmospheric Physics.

Received 14 February 1981.
0003-6935/81/172925-04$00.50/0.

data by a straight line will give more weight to data points at large air mass (i.e., low sun) than to data at small air mass.

Although the expressions for a weighted least-squares fit may be found in a great many sources, we choose to include our own derivation in Sec. II, as we believe it provides greater insight, and to show a proper justification for its use. In comparing the various weighted and unweighted expressions in Sec. III, we will assume that we know the relevant parameters of the underlying population of optical thicknesses. Although we can never know these values from a small sample, this assumption is quite valid in comparing the accuracy of different statistical expressions.

II. Proposed Technique: The Weighted Fit

We here propose a technique which will weight all data points equally, a physically more plausible technique with the assumption of purely random variations in τ during the period of the measurements. To accomplish this, we first define an instantaneous optical depth τ_n by

$$\tau_n = \frac{\ln F_0 - \ln F_n}{m_n} \tag{4}$$

and a mean optical detph $\bar{\tau}$ by

$$\bar{\tau} = \frac{1}{N} \sum_{n=1}^{N} \tau_n, \tag{5}$$

where the subscript n indicates a particular measurement, and N is the total number of measurements in the sample. The fluctuation in optical depth $\Delta\tau_n$ from the mean for the day for any measurement F_n is then simply $\Delta\tau_n = \tau_n - \bar{\tau}$.

Ideally, the value $\ln F_0$ appearing in Eq. (4) is the exact value which, of course, is unknown. It thus becomes an object of the data analysis to determine a best estimate of $\ln F_0$, $\ln F_0^*$, before values for τ_n and $\bar{\tau}$ may be obtained. In the present method, it is assumed that variations in τ throughout the day are random, and we thus seek the value of the intercept $\ln F_0^*$ so that

$$\sum_{n-1} (\tau_n - \bar{\tau})^2$$

is minimized. This is equivalent to finding the value $\ln F_0^*$ so that

$$\frac{d}{d \ln F_0^*} \sum_{n=1}^{N} (\tau_n - \bar{\tau})^2 = 0, \tag{6}$$

or upon taking the derivative

$$2 \sum_{n=1}^{N} (\tau_n - \bar{\tau}) \left(\frac{d\tau_n}{d \ln F_0^*} - \frac{d\bar{\tau}}{d \ln F_0^*} \right) = 0. \tag{7}$$

Using the relationships

$$\frac{d\tau_n}{d \ln F_0^*} = \frac{1}{m_n} \tag{8}$$

$$\frac{d\bar{\tau}}{d \ln F_0^*} = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{m_n}, \tag{9}$$

Eq. (7) may be shown to yield the following expression for $\ln F_0^*$:

$$\ln F_0^* = \left(N \sum_{n=1}^{N} \ln F_n m_n^{-2} - S_{-1} \sum_{n=1}^{N} \ln F_n m_n^{-1} \right) \bigg/ \Delta, \tag{10}$$

where

$$\Delta = NS_{-2} - S_{-1}^2, \tag{11}$$

and we have introduced the notation

$$S_j = \sum_{n=1}^{N} m_n^j. \tag{12}$$

When this result is inserted into Eqs. (4) and (5), an equivalent expression for $\bar{\tau}^*$ may be obtained:

$$\bar{\tau}^* = \left(S_{-1} \sum_{n=1}^{N} \ln F_n m_n^{-2} - S_{-2} \sum_{n=1}^{N} \ln F_n m_n^{-1} \right) \bigg/ \Delta. \tag{13}$$

Equation (10) thus yields the value of the intercept $\ln F_0^*$, which, with the calculated mean optical depth $\bar{\tau}$, minimizes the deviation of the instantaneous optical depth about the mean value in a least-squares sense. This procedure is reasonable with the assumption that the values of τ vary randomly throughout the period of measurement, and it gives equal weighting to each measurement. The standard procedure of least-squares fit to the log of the measured solar fluxes $\ln F$ inherently weights the variations of optical depth $\Delta\tau$ by the air mass m as we have shown [i.e., Eq. (3)] thereby weighting measurements at low solar angles more than those at high sun. The present procedure should thus provide better values for both the optical depth and the intercept (i.e., solar constant) on days when there is not a systematic temporal change in τ occurring. On these latter days neither method will provide reliable results, and one must resort to calculating instantaneous values of τ from each individual measurement, providing a reliable value of the intercept has been established. On the other hand, on very stable days the variations in τ are very slight, and both methods will probably produce nearly identical results.

In a later paper, experimental testing of the present method and comparison with the older technique will be presented.

III. Statistical Analysis

In this Section, we propose to examine the formulas given by Eqs. (10) and (13) from a statistical point of view and to compare them with the old unweighted formulas, which we now quote for reference (we shall use a subscript u to designate all unweighted results):

$$\ln F_{0u}^* = \left(S_2 \sum_{n=1}^{N} \ln F_n - S_1 \sum_{n=1}^{N} \ln F_n m_n \right) \bigg/ \Delta_u; \tag{14}$$

$$\bar{\tau}_u^* = -\left(N \sum_{n=1}^{N} \ln F_n m_n - S_1 \sum_{n=1}^{N} \ln F_n \right) \bigg/ \Delta_u; \tag{15}$$

$$\Delta_u = NS_2 - S_1^2. \tag{16}$$

We now assume that as measurements are made throughout the day the real instantaneous optical depths τ are drawn from a population with a mean of τ_0 and with a standard deviation $\delta\tau$ caused by real atmospheric fluctuations. Our measurements then represent a sample drawn from that population.

The first result which we seek is the expectation value of the two parameters we have obtained. From Eq. (2), we have

$$\langle \ln F_n \rangle = \ln F_0 - m_n \tau_0, \quad (17)$$

which may be inserted into Eqs. (10)–(16) to show that

$$\langle \ln F_0^* \rangle = \langle \ln F_{0u}^* \rangle = \ln F_0, \quad (18)$$

$$\langle \bar{\tau}^* \rangle = \langle \bar{\tau}_u^* \rangle = \tau_0, \quad (19)$$

i.e., neither formula exhibits systematic errors.

That being the case, we now turn our attention to expected deviations from the mean (or true) values. The expressions for the standard deviations of $\ln F_0^*$ and $\bar{\tau}^*$ in the weighted case may be taken straight out of the standard references[4,5]:

$$\sigma^2(\ln F_0^*) = N\delta\tau^2/\Delta \quad (20)$$

$$\sigma^2(\bar{\tau}^*) = S_{-2}\delta\tau^2/\Delta, \quad (21)$$

where Δ is again given by Eq. (11), and we write $\delta\tau^2$ for $(\delta\tau)^2$.

The expressions in the unweighted case cannot be so easily obtained, and we must resort to a longer derivation (outlined in the Appendix) to arrive at the results:

$$\sigma^2(\ln F_{0u}^*) = \delta\tau^2(S_2^3 - 2S_1S_2S_3 + S_1^2S_4)/\Delta_u^2 \quad (22)$$

$$\sigma^2(\bar{\tau}_u^*) = \delta\tau^2(N^2S_4 - 2NS_1S_3 + S_1^2S_2)/\Delta_u^2. \quad (23)$$

Equations (20)–(23) are rather complicated and it is not possible to decide by inspection which form is the smaller or by how much. We have, therefore, performed some simple calculations with these equations. We have taken a series of cases with N ranging from 3 to 20 and air-mass values starting at $m = 1$ and increasing in steps of 0.5 up to $m = (N + 1)/2$. In Table I we list the values of the standard deviation (i.e., the square root of the variance) divided by $\delta\tau$.

It can be seen that for small N, when all the air-mass values are small, there is little to choose between the two formulations, although we may note that the new (weighted) expressions are always superior. As N increases and the number of large air-mass values increases with it, the superiority of the new formulations increases steadily. Although the selection of air-mass values employed in this calculation is not typical of most radiometry programs, the results presented in Table I make it clear that, as more readings are made at large air-mass values, the advantages of the weighted expressions become more and more apparent. This, of course, is precisely what we would have expected on the basis of our earlier discussion.

Table I. Renormalized Standard Deviations for the Weighted and Unweighted Cases

		Weighted		Unweighted	
N	$m_{\max}$	$\sigma(\ln F_0^*)/\delta\tau$	$\sigma(\bar{\tau}^*)/\delta\tau$	$\sigma(\ln F_{0u}^*)/\delta\tau$	$\sigma(\bar{\tau}_u^*)/\delta\tau$
3	2.0	2.777	2.087	3.009	2.236
4	2.5	2.195	1.495	2.557	1.691
5	3.0	1.878	1.177	2.337	1.393
6	3.5	1.677	0.979	2.216	1.202
7	4.0	1.537	0.844	2.146	1.069
8	4.5	1.434	0.745	2.107	0.970
9	5.0	1.354	0.669	2.087	0.893
10	5.5	1.291	0.610	2.079	0.830
11	6.0	1.239	0.562	2.080	0.799
12	6.5	1.196	0.522	2.089	0.736
13	7.0	1.159	0.488	2.098	0.699
14	7.5	1.128	0.459	2.112	0.667
15	8.0	1.100	0.434	2.129	0.639
16	8.5	1.076	0.413	2.148	0.614
17	9.0	1.055	0.393	2.169	0.592
18	9.5	1.036	0.376	2.190	0.572
19	10.0	1.019	0.361	2.212	0.554
20	10.5	1.003	0.347	2.235	0.537

IV. Conclusion

We see from Table I and he accompanying discussion that although neither the weighted nor the unweighted formulas lead to any systematic errors in the estimation of solar constant and average optical depth, the unweighted formulas are subject to greater statistical uncertainties as evidenced by their higher variances. Thus we may conclude that, on average, use of the new weighted formulas will yield results which are closer to the true values of these parameters. This is especially true in the case of the solar constant, indicating that the new formulas should definitely be used in measurement programs designed to monitor this most important quantity.

This project was funded by contract N00014-76-C-0438.

Appendix

Here we sketch the steps needed to arrive to Eqs. (22) and (23). From the definition of the variance, we have

$$\begin{aligned}
\sigma^2(\ln F_{0u}^*) &= \langle \ln^2 F_{0u}^* \rangle - \langle \ln F_{0u}^* \rangle^2 \\
&= \left\langle S_2^2\left(N \ln F_{0u}^* - \sum_{n=1}^{N} \tau_n m_n\right)^2 - 2S_1S_2 \right. \\
&\quad * \left(N \ln F_{0u}^* - \sum_{n=1}^{N} \tau_n m_n\right)\left(\ln F_{0u}^* S_1 - \sum_{n=1}^{N} \tau_n m_n^2\right) \\
&\quad \left. + S_1^2 \left(\left(\ln F_{0u}^* S_1 - \sum_{n=1}^{N} \tau_n m_n^2\right)^2\right) \right\rangle \Big/ \Delta_u^2 - \ln^2 F_0 \\
&= \left[-2N \ln F_0 S_2^2 \left\langle \sum_{n=1}^{N} \tau_n m_n \right\rangle + S_2^2 \left\langle \left(\sum_{n=1}^{N} \tau_n m_n\right)^2 \right\rangle \right. \\
&\quad + 2N \ln F_0 S_1 S_2 \left\langle \sum_{n=1}^{N} \tau_n m_n^2 \right\rangle \\
&\quad + 2 \ln F_0 S_2 S_1^2 \left\langle \sum_{n=1}^{N} \tau_n m_n \right\rangle \\
&\quad - 2S_1S_2 \left\langle \sum_{n=1}^{N} \tau_n m_n \sum_{n=1}^{N} \tau_n m_n^2 \right\rangle + S_1^2 \left\langle \left(\sum_{n=1}^{N} \tau_n m_n^2\right)^2 \right\rangle \\
&\quad \left. - 2 \ln F_0 S_1^3 \left\langle \sum_{n=1}^{N} \tau_n m_n^2 \right\rangle \right] \Big/ \Delta_u^2 \\
&= \delta\tau^2(S_2^3 - 2S_1S_2S_3 + S_1^2S_4)/\Delta_u^2,
\end{aligned}$$

and we have made use of the following results:

$$\left\langle \sum_{n=1}^{N} \tau_n m_n^j \right\rangle = \tau_0 \sum_{n=1}^{N} m_n^j = \tau_0 S_j;$$

$$\left\langle \left(\sum_{n=1}^{N} \tau_n m_n \right)^2 \right\rangle = \delta\tau^2 \sum_{n=1}^{N} m_n^2 + \tau_0^2 \left(\sum_{n=1}^{N} m_n \right)^2 = \delta\tau^2 S_2 + \tau_0^2 S_1^2;$$

$$\left\langle \left(\sum_{n=1}^{N} \tau_n m_n^2 \right)^2 \right\rangle = \delta\tau^2 \sum_{n=1}^{N} m_n^4 + \tau_0^2 \left(\sum_{n=1}^{N} m_n^2 \right)^2 = \delta\tau^2 S_4 + \tau_0^2 S_2^2;$$

$$\left\langle \sum_{n=1}^{N} \tau_n m_n \sum_{n=1}^{N} \tau_n m_n^2 \right\rangle = \delta\tau^2 \sum_{n=1}^{N} m_n^3 + \tau_0^2 \sum_{n=1}^{N} m_n \sum_{n=1}^{N} m_n^2 = \delta\tau^2 S_3 + \tau_0^2 S_1 S_2.$$

Equation (23) for $\sigma^2(\overline{\tau}_0^*)$ may be obtained in an analogous manner.

References

1. C. G. Abbot and L. B. Aldrich, *Annals of the Astrophysical Observatory of the Smithsonian Institutioin*, Vol. 3 (U.S. GPO, Washington, D.C., 1913).
2. A. Ångstrom, Tellus **13**, 214 (1961).
3. G. E. Shaw, J. A. Reagan, and B. M. Herman, J. Appl. Meteorol. **12**, 374 (1973).
4. S. L. Meyer, *Data Analysis for Scientists and Engineers* (Wiley, New York, 1975).
5. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1969).

Meetings Calendar *continued from page 2924*

1982

January

18–21 SPIE 1st Int. Conf. on Distributed Computerized Picture Information Systems, Newport Beach *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

25–29 SPIE Los Angeles Technical Symp. & Exhibit, Los Angeles *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

February

4–9 Plasma Spectrochemistry Conf., Orlando *1982 Winter Conf., c/o ICP Inf. Newsletter Chem., GRC Towers, U. Mass., Amherst, Mass. 01003*

March

3 FIPS Software Documentation Workshop, Gaithersburg *Documentation Workshop, NBS, Inst. Computer Sciences & Tech., Bldg. 225, Rm. A-265, Wash., D.C. 20234*

8–10 Laser Techniques For Extreme Ultraviolet Spectroscopy, OSA Top. Mtg., Boulder *Mtgs. Dept., OSA, 1816 Jefferson Place, N.W., Wash., D.C. 20036*

8–11 SPIE Instrumentation in Astronomy IV Conf., Tucson *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

8–12 Pittsburgh Conf. & Expo. on Analytical Chemistry & Applied Spectroscopy, Atlantic City *M. Senneway, 405 Carmel Dr., Aliquippa, Pa. 15001*

14–18 6th Symp. of Temperature—Its Measurement and Control in Science and Industry, Wash., D.C. *J. Schooley, NBS, Rm. B128, Phys. Bldg., Wash., D.C. 20234*

15–19 Photographic Science course, Rochester *V. Johnson, RIT-GARC, 1 Lomb Dr., Rochester, N.Y. 14623*

29–1 Apr. SPIE San Jose Semiconductor Confs., San Jose *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

29–2 Apr. 183rd ACS Natl. Mtg., Las Vegas *A. T. Winstead, 1155 16th St. N.W., Wash., D.C. 20036*

April

? SPIE 1st Int. Conf. on New Technology Telescopes *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

12–16 SPIE Int. Tech. Symp. East, Wash., D.C. *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

13–15 **Optical Fiber Communication, 5th OSA Top. Mtg.,** Phoenix *Mtgs. Dept., OSA, 1816 Jefferson Pl., N.W., Wash., D.C. 20036*

14–16 **Lasers and Electrooptics Conf.,** Phoenix *Mtgs. Dept., OSA, 1816 Jefferson Pl. N.W., Wash., D.C. 20036*

May

9–11 SPIE Application of Optical Instrumentation in Medicine X Conf., New Orleans *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

17–21 **Applied Optics 1982 Spring Conf.,** Rochester Mtgs. Dept. *OSA, 1816 Jefferson Pl., N.W., Wash., D.C. 20036*

June

22–25 OSA XII Int. Quantum Electronics Conf., Munich *Mtgs. Dept., OSA, 1816 Jefferson Place, N.W., Wash., D.C. 20036*

28–1 July Precision Electromagnetic Measurements Conf., Boulder *D. Belsher, NBS, 1-4001, 325 Broadway, Boulder, Colo. 80303*

July

5–7 Applications of Laser-Doppler Anemometry to Fluid Mechanics Int. Symp., Lisbon *F. Durst, Sonderforschungsbereich 80 An Der Universitat Karlsruhe, 7500 Karlsruhe 1, Postfach 6380, FRG*

August

21–27 SPIE 26th Ann. Int. Tech. Symp. & Exhibit/15th Int. Cong. on High Speed Photography and Photonics, San Diego *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

23–27 X-Ray & Atomic Inner-Shell Physics Int. Conf., Eugene, Ore. *B. Crasemann, Phys. Dept., U. Ore., Eugene, Ore. 97403*

September

? SPIE Industrial Applications of Infrared Thermography Conf., Milwaukee *SPIE, P.O. Box 10, Bellingham, Wash. 98227*

5–10 Precision and Speed in Close Range Photogrammetry Int. Symp., York *K. Atkinson, Dept. Photogrammatry & Surveying, University Coll. London, Gower St., London WC1E 6BT, England*

continued on page 2945