NOAA Technical Report NESDIS 123

Comparing Two Sets of Noisy Measurements

Washington, D.C.
April 2007

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service

NOAA TECHNICAL REPORTS

National Environmental Satellite, Data, and Information Service

The National Environmental Satellite, Data, and Information Service (NESDIS) manages the Nation's civil Earth-observing satellite systems, as well as global national data bases for meteorology, oceanography, geophysics, and solar-terrestrial sciences. From these sources, it develops and disseminates environmental data and information products critical to the protection of life and property, national defense, the national economy, energy development and distribution, global food supplies, and the development of natural resources.

Publication in the NOAA Technical Report series does not preclude later publication in scientific journals in expanded or modified form. The NESDIS series of NOAA Technical Reports is a continuation of the former NESS and EDIS series of NOAA Technical Reports and the NESC and EDS series of Environmental Science Services Administration (ESSA) Technical Reports.

An electronic copy of this technical report (NESDIS 123) may be obtained at http://www.orbit.nesdis.noaa.gov/smcd/spb/ozone/pubs_docs.html

A limited number of copies of earlier reports are available by contacting Jessica Pejsa, NOAA/NESDIS, E/RA3, 5200 Auth Road, Room 601, Camp Springs, Maryland 20746, (301) 763-8184 x115. Copies can also be ordered from the National Technical Information Service (NTIS), U.S. Department of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, VA 22161, (703) 487-4650 (prices on request for paper copies or microfiche, please refer to PB number when ordering). A partial listing of more recent reports appears below:

NESDIS 95 Preliminary Findings from the Geostationary Interferometer Observing System Simulation Experiments (OSSE). Bob Aune, Paul Menzel, Jonathan Thom, Gail Bayler, Allen Huang, and Paolo Antonelli, June 2000.

NESDIS 96 Hydrography of the Ross Sea Continental Shelf During the Roaverrs, NBP96-06, Cruise December 1996 - January 1997. Michael L. Van Woert, David Pryor, Eric Quiroz, Richard Slonaker, and William Stone, September 2000.

NESDIS 97 Hydrography of the Ross Sea Continental Shelf During the Roaverrs, NBP97-09, Cruise December 1997 - January 1998. Michael L. Van Woert, Lou Gordon, Jackie Grebmeier, Randal Holmbeck, Thomas Henderson, and William F. Van Woert, September 2000.

NESDIS 98 NOAA-L and NOAA-M AMSU-A Antenna Pattern Corrections. Tsan Mo, August 2000.

NESDIS 99 The Use of Water Vapor for Detecting Environments that Lead to Convectively Produced Heavy Precipitation and Flash Floods. Rod Scofield, Gilberto Vicente, and Mike Hodges, September 2000.

NESDIS 100 The Resolving Power of a Single Exact-Repeat Altimetric Satellite or a Coordinated Constellation of Satellites: The Definitive Answer and Data Compression. Chang-Kou Tai, April 2001.

NESDIS 101 Evolution of the Weather Satellite Program in the U.S. Department of Commerce - A Brief Outline. P. Krishna Rao, July 2001.

NESDIS 102 NOAA Operational Sounding Products From Advanced-TOVS Polar Orbiting Environmental Satellites. Anthony L. Reale, August 2001.

NESDIS 103 GOES-11 Imager and Sounder Radiance and Product Validations for the GOES-11 Science Test. Jaime M. Daniels and Timothy J. Schmit, August 2001.

NOAA Technical Report NESDIS 123

Comparing Two Sets of Noisy Measurements

Lawrence E. Flynn
NOAA/NESDIS/STAR
5200 Auth Road
Camp Springs, MD 20746

Washington, DC
April 2007

U.S. DEPARTMENT OF COMMERCE
Carlos M. Gutierrez, Secretary

National Oceanic and Atmospheric Administration
Vice Admiral Conrad C. Lautenbacher, Jr., U.S. Navy (Ret.), Under Secretary

National Environmental Satellite, Data, and Information Service
Mary Kicza, Assistant Administrator

TABLE OF CONTENTS

TABLES

FIGURES

PREFACE

In validation studies, comparisons frequently arise in which one has two sets of noisy measurements of the same quantity, often obtained by some match-up process. The results are often presented on a scatter plot with a reference line. This Technical Report discusses choices for such reference lines and attempts to provide some guidance on the use of standard linear fits and other calculations to learn more about the relationship between the two sets of data. The first section uses a simple example to provide a setting for a discussion of the theoretical problem. The second section uses more creative synthetic data sets to expose some of the difficulties. The third and fourth sections present some real world applications and further discussion of the complications. The fifth section summarizes the results, and the sixth section provides an outline with a basic set of calculations, analysis, questions and plots as a guide for investigating such data.

PROBLEM STATEMENT

We define an experiment where one can measure a quantity with two separate methods, A and B, where both methods provide noisy estimates, linearly related to the true values. Specifically, let $\{R_i\}$ be a set of truth values, and let $\{A_i\}$ and $\{B_i\}$ be two sets of measurements (or retrievals from measurements) with models and measurements of the form:

$$A = \mu_A R + a \quad \text{and} \quad A_i = \mu_A R_i + a + \varepsilon_i$$

and

$$B = \mu_B R + b \quad \text{and} \quad B_i = \mu_B R_i + b + \delta_i$$

where μ_A and μ_B are slopes, a and b are constants, and ε_i and δ_i are measurement noise contributions. Eliminating R in the models and measurements gives

$$A = \mu_A/\mu_B\, B + c \quad \text{and} \quad A_i = \mu_A/\mu_B\, (B_i - \delta_i) + c + \varepsilon_i$$
$$A = \mu_C\, B + c \quad \text{and} \quad A_i = \mu_C\, (B_i - \delta_i) + c + \varepsilon_i$$

or

$$B = \mu_B/\mu_A\, A + d \quad \text{and} \quad B_i = \mu_B/\mu_A\, (A_i - \varepsilon_i) + d + \delta_i$$
$$B = \mu_D\, A + d \quad \text{and} \quad B_i = \mu_D\, (A_i - \varepsilon_i) + d + \delta_i$$
$$B = 1/\mu_C\, A + d \quad \text{and} \quad B_i = 1/\mu_C\, (A_i - \varepsilon_i) + d + \delta_i$$

Where the slopes are now $\mu_C=\mu_A/\mu_B$ in the first case and its reciprocal, $\mu_D=1/\mu_C=\mu_B/\mu_A$, in the second, and c and d are constants algebraically related to μ_A, μ_B, a, and b.

The question addressed in this report is: Given two sets of measurements, how well can we estimate the parameters in these models when $\{R_i\}$ is unknown? The basic problem is that we are trying to estimate characteristics of the linear relationship between A and B (for example, the slope and one point) and the variances of $\{\varepsilon_i\}$ and $\{\delta_i\}$ from only five quantities – the means and variances of $\{A_i\}$ and $\{B_i\}$ and the covariance of $\{A_i\}$ with $\{B_i\}$. We will see how far standard linear regression can take us, and explore the impact of three different simplifying assumptions. The individual simplifying assumptions are as follows: 1) The two sets of measurements have the same slopes with respect to the truth, that is, $\mu_A/\mu_B = 1$; 2) One of the measurements is known to have no or very little noise, *e.g.*, $\delta_i \equiv 0$ or the δ_i are very small; or 3) The noise in the two measurements, $\{\varepsilon_i\}$ and $\{\delta_i\}$, are of similar relative sizes/variances.

SECTION 1. Warm-up Mathematical Exercises

For simplicity of exposition, we create a small truth data set of twelve values and two very non-random sets of "measurements" with $\mu_A = \mu_B = 1$, and $a = b = 0$.

The truth values are given in the first row of Table 1, and the means and variances for these sets appear in the last two columns. The $\{A_i\}$ and $\{B_i\}$ are created so that the differences for A_i and B_i values from the truth, R_i are uncorrelated (have zero covariance) and have zero mean, and so that the A_i data is "noisier" than the B_i data. The data have been "cooked" so well that they have zero mean differences for each of the three possible truth values.

Table 1. Very Simple Illustrative Example.

I	1	2	3	4	5	6	7	8	9	10	11	12	Mean	Variance
R_i	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	2.0	2.0	2.0	2.0	1.0	$0.\overline{66}\overline{6}$=2/3
A_i	-0.4	-0.4	0.4	0.4	0.6	0.6	1.4	1.4	1.6	1.6	2.4	2.4	1.0	$0.82\overline{6}$=2/3+0.16
B_i	-0.2	0.2	-0.2	0.2	0.8	1.2	0.8	1.2	1.8	2.2	1.8	2.2	1.0	$0.70\overline{6}$=2/3+0.04
$\varepsilon_i=A_i-R_i$	-0.4	-0.4	0.4	0.4	-0.4	-0.4	0.4	0.4	-0.4	-0.4	0.4	0.4	0.0	0.160=0.4^2
$\delta_i=B_i-R_i$	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	-0.2	0.2	0.0	0.040=0.2^2
A_i-B_i	-0.2	-0.6	0.6	0.2	-0.2	-0.6	0.6	0.2	-0.2	-0.6	0.6	0.2	0.0	0.200=0.16+0.04

Notice that

$A = R$ and $A_i = R_i + \varepsilon_i$, with ε_i = {-0.4,-0.4,0.4,0.4,-0.4,-0.4,0.4,0.4,-0.4,-0.4,0.4,0.4}

$B = R$ and $B_i = R_i + \delta_i$, with δ_i = {-0.2,0.2,-0.2,0.2,-0.2,0.2,-0.2,0.2,-0.2,0.2,-0.2,0.2}

$\overline{R}_i = \overline{A}_i = \overline{B}_i = 1$, $\overline{A}_i - \overline{R}_i = \overline{B}_i - \overline{R}_i = \overline{A}_i - \overline{B}_i = 0$, and $\overline{\varepsilon}_i = \overline{\delta}_i = 0$

and the covariance of $\{A_i\}$ with $\{B_i\}$ is

$COV(A_i,B_i) = COV(B_i,A_i) = 2/\overline{3} = VAR(R_i)$

The $\overline{\ }$ operator is used to denote the mean of a set of values. The mathematical symbol $\overline{\ }$ is used to denote a repeating decimal. The VAR() operator is used to denote the variance of a set of values and the COV(,) operator is used to denote the covariance of two sets of values. (Aside: If these were real samples of random variables, and the means were estimated from the samples, one would compute the variance estimates with the estimated mean and a denominator of n–1=11 to obtain an unbiased estimate of the variances for the underlying distributions of A and B, instead of a denominator of n=12 as was used here.)

If we ignore the fact that both measurements are noisy, and proceed with standard linear regression fits of $\{A_i\}$ with $\{B_i\}$ and $\{B_i\}$ with $\{A_i\}$, then we find the following:

1) For $\{A_i\}$ with $\{B_i\}$, the fit line passes through the point $(\overline{B}_i,\overline{A}_i)=(1,1)$ and the fit slope estimate is

$m_{AB} = m_C = COV(A_i,B_i)/VAR(B_i) = 2/\overline{3} / (2/\overline{3} + 0.04) \approx 0.943 < 1 = \mu_C = \mu_A/\mu_B$

2) For $\{B_i\}$ with $\{A_i\}$, the fit line passes through the point $(\overline{A}_i,\overline{B}_i)=(1,1)$ and the fit slope estimate is

$m_{BA} = m_D = COV(B_i,A_i)/VAR(A_i) = 2/\overline{3} / (2/\overline{3} + 0.16) \approx 0.806 < 1 = \mu_D = 1/\mu_C = \mu_B/\mu_A$

Both slope estimates are smaller than the true model slopes, and the one found using the noisier measurement as the independent variable is poorer. The problem is that standard linear regression methods expect the independent variable to be noise-free. The variances in both of the denominators for these two slope estimates have additional contributions from the noise in the independent variable choice.

A scatter plot of B versus A appears in Fig. 1 along with two fit lines and a reference line. The data are shown by the + symbols. The ✳ symbols are binned values discussed in Subsection 1.3. The solid reference line is just B=A. The dotted line is from a linear regression fit of the B values with the A values. It has a B/A slope of 0.806 and a B-intercept of 0.194. The dashed line is from a linear regression fit of the A values with the B values. It has an A/B slope of 0.943 (B/A slope of 1/0.943≈1.060) and an A-intercept of 0.057. Notice that both fits have slopes less than 1.0 in their appropriate coordinate systems, pass through the point (1,1), and that the fit using the noisier A data as the independent variable has the poorest slope estimate. The slope estimate from the linear regression fit of the B values with A values is equal to $COV(A_i,B_i)/VAR(A_i)$ and from the linear regression fit of the A values with B values is equal to $COV(A_i,B_i)/VAR(B_i)$. In the next subsection, we consider the problem of estimating the truth and measurement noise variability given two match-up data sets.

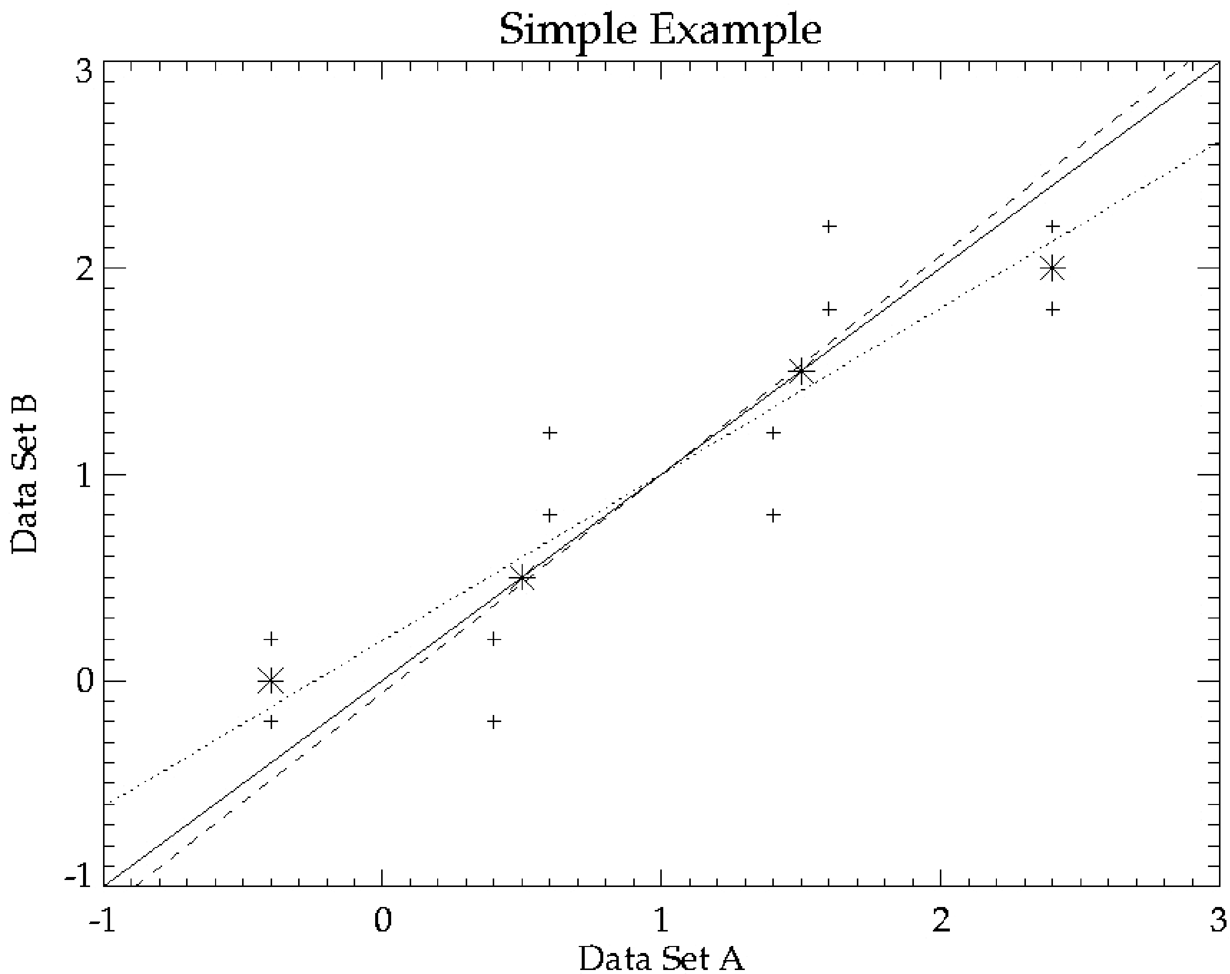


Figure 1. Scatter plot of B versus A for the simple example in Table 1 along with two fit lines and a reference line. The data are shown by the + symbols. The solid reference line is just A=B. The dotted line is from a linear regression fit of the B values with the A values. It has a B/A-slope of 0.806 and a B-intercept of 0.194. The dashed line is from a linear regression fit of the A values with the B values. It has an A/B-slope of 0.943 and an A-intercept of 0.057. Notice that both fits have slopes less than 1.0 in their appropriate coordinates, pass through the point (1,1), and that the fit using A has the smaller slope. The ✳ symbols are the results of binning for each unit interval using the A data to determine the bins.

Subsection 1.1. Estimating Measurement Noise and Truth Variability

If we have two sets of independent, noisy measurements of the same set of truth values, can we use them to estimate their own noise variability? The answer is a qualified "Yes" and one possible qualification is that $\mu_C = \mu_A/\mu_B$ is known *a priori*. (As we will see later, other possibilities involve some knowledge of the noise or signal variability.) We also make the standard assumptions that both measurements have independent, zero mean measurement noise, that is,

$$A_i = \mu_A R_i + a + \varepsilon_i$$

and

$$B_i = \mu_B R_i + b + \delta_i$$

where ε_i and δ_i are instances of independent random variables with zero mean, uncorrelated with each other and with the truth, and variances σ_ε^2 and σ_δ^2 respectively. Substituting these expressions we find that

$$VAR(A_i) = \mu_A^2\, VAR(R_i) + 2{*}\mu_A COV(R_i,\varepsilon_i) + VAR(\varepsilon_i) \approx \mu_A^2\, VAR(R_i) + \sigma_\varepsilon^2$$
$$VAR(B_i) = \mu_B^2\, VAR(R_i) + 2{*}\mu_B COV(R_i,\delta_i) + VAR(\delta_i) \approx \mu_B^2\, VAR(R_i) + \sigma_\delta^2$$
$$COV(A_i,B_i) = \mu_A\mu_B VAR(R_i) + \mu_B COV(R_i,\varepsilon_i) + \mu_A COV(R_i,\delta_i) + COV(\varepsilon_i,\delta_i) \approx \mu_A\mu_B VAR(R_i)$$

(The lack of correlation of the measurement noise distributions with other quantities means that we expect terms such as $COV(R_i,\delta_i)$ and $COV(\varepsilon_i,\delta_i)$ to be unbiased and small, and have dropped them in deriving these approximate relationships. Their contributions are important for obtaining confidence intervals for slope estimates.) If $\mu_A/\mu_B = \mu_C$ is known, then these can be manipulated to give:

For the general case,	For our simple example, $\mu_A=\mu_B=1$ and $\mu_C=\mu_A/\mu_B=1$. These become
$\mu_B\mu_A VAR(R_i) \approx COV(A,B)$	$VAR(R_i) \approx 2/3$
$\sigma_\varepsilon^2 \approx VAR(A_i) - COV(A,B) * \mu_C$	$\sigma_\varepsilon^2 \approx (2/3 + 0.16) - 2/3 * 1 = 0.16$
$\sigma_\delta^2 \approx VAR(B_i) - COV(A,B) / \mu_C$	$\sigma_\delta^2 \approx (2/3 + 0.04) - 2/3 / 1 = 0.04$

and we can estimate two of the three values directly from the sets A and B without knowing R. Practical difficulties are examined in Sections 2 and 3. Simple offset biases in the A and B measurements would not affect these calculations. (A common assumption is that $\mu_A = \mu_B$, that is, $\mu_C = 1$. If this is incorrectly assumed, and we proceed with the calculations, then the approach breaks down because we do not completely cancel the $\mu_B\mu_A VAR(R_i)$ terms in finding estimates of the noise variability. We would underestimate the noise in the data set with the smaller slope versus the truth and overestimate it in the one with the larger slope. This occurs in the real world examples in Section 3.) Lastly, no matter what the actual value of $\mu_A/\mu_B = \mu_C$ is, the variance of the difference

$$VAR(A_i - B_i) = VAR(A_i) + VAR(B_i) - 2*COV(A_i,B_i) \approx (\mu_A - \mu_B)^2 VAR(R_i) + VAR(\varepsilon_i) + VAR(\delta_i)$$
$$\approx (\mu_A - \mu_B)^2 VAR(R_i) + \sigma_\varepsilon^2 + \sigma_\delta^2 = \mu_B^2(\mu_C - 1)^2 VAR(R_i) + \sigma_\varepsilon^2 + \sigma_\delta^2$$

gives an approximate upper bound on the sum of the measurement noise variances, $\sigma_\varepsilon^2 + \sigma_\delta^2$, and this estimate is better when the difference between μ_A and μ_B is smaller, *i.e.*, the closer μ_C is to 1.

Subsection 1.2. Adjusting the Slope Estimates for Measurement Noise in the Independent Variable

This subsection answers the question: Why are both slopes less than 1 in the fits for the simple example? Standard linear least squares regression calculations (See, e.g., Neter and Wasserman, 1974.) assume that there is no noise or error in the independent variable, that is, there is a model with

$$A = \mu_C C + c$$

and a set of noisy measurements of A with

$$A_i = \mu_C C_i + c + \varepsilon_i$$

The slope estimate for $\{A_i\}$ fit with $\{C_i\}$ is given by

$$m_C = COV(A_i,C_i)/VAR(C_i)$$

Replacing A_i with its expression in terms of C_i gives

$$m_C = \mu_C + COV(\varepsilon_i,C_i)/VAR(C_i) \approx \mu_C$$

This final form is of no help in computing m_C, since μ_C is the unknown, but, practically, it does tell us that having either less measurement noise (smaller absolute values of ε_i) or a more variable range of observations [larger $VAR(C_i)$] will lead to better regression estimates, m_C, of the slope, μ_C, and that incidental covariance between the measurement noise and independent variable [larger $COV(\varepsilon_i,C_i)$] leads to poorer estimates. Even though there is a plus sign in front of the $COV(\varepsilon_i,C_i)$ term in the expression for m_C, there is no reason to expect the covariance involved to be positive rather than negative. So, with the standard assumptions, m_C will provide an unbiased estimate of μ_C.

What happens if we use a noisy set of measurements $\{B_i\}$, instead of $\{C_i\}$, with a model of the form

$$A = \mu_A/\mu_B\, B + c$$

and a set of noisy measurements of A with

$$A_i = \mu_A/\mu_B\,(B_i - \delta_i) + c + \varepsilon_i$$

in our comparisons? (In standard statistics texts, this class of problems is referred to as "fits with errors in both coordinates." See Press, 1992, Section 15.3 for methods to handle this complication directly. This subsection investigates what happens when standard linear regression meets this problem.) If we knew the δ_i's, then we could find $C_i = B_i - \delta_i$ and be back to the first case. If we do not know them, and proceed by ignoring the presence of the δ_i noise term, then the slope estimate for $\{A_i\}$ fit with $\{B_i\}$ is given by

$$m_{AB} = COV(B_i,A_i)/VAR(B_i)$$

Substituting for A_i in terms of B_i gives

$$m_{AB} = \mu_A/\mu_B\,[1 - COV(B_i,\delta_i)/VAR(B_i)] + COV(B_i,\varepsilon_i)/VAR(B_i)$$

We also have $COV(B_i,\delta_i) = \mu_B COV(R_i,\delta_i) + VAR(\delta_i)$, and if we ignore the $COV(B_i,\varepsilon_i)$ and $COV(R_i,\delta_i)$ terms (because we expect them to be small and because we expect them to be unbiased), we find

$$m_{AB} = COV(A_i,B_i)/VAR(B_i) \approx \mu_A/\mu_B\,[1 - VAR(\delta_i)/VAR(B_i)]$$

Variances are always non-negative, *so the slope estimate, m_{AB}, tends to be less in absolute value than the true slope, μ_A/μ_B*. (The goal in least squares regression is to minimize the sum of the residuals squared. If there is noise in the independent variable, then a larger magnitude slope will increase this component in the residual difference, thus the slope is diminished in magnitude from the result obtained for a non-noisy measurement of the independent variable.) Notice that this error gets worse for larger noise in the independent variable (larger relative to the variability).

We could have just as easily fit B with A and obtained an estimate for the reciprocal slope,

$$m_{BA} = COV(A_i,B_i)/VAR(A_i) \approx \mu_B/\mu_A\,[1 - VAR(\varepsilon_i)/VAR(A_i)]$$

So the slope estimate, m_{AB}, of μ_A/μ_B from fitting $\{A_i\}$ with $\{B_i\}$ tends to be less in absolute value than the actual slope, and the slope estimate, m_{BA}, of its reciprocal, μ_B/μ_A, from fitting $\{B_i\}$ with $\{A_i\}$ also tends to be less in absolute value than the actual slope. This explains why both slopes for the fits of the simple data sets were less than the model slope ratios of 1.

If we have estimates of the variability of the measurement noise for one of the sets, then we can invert the approximate expression for m_{AB} to obtain an approximately unbiased estimate of μ_A/μ_B as follows:

$$\mu_A/\mu_B \approx m_{AB} / [1 - \sigma_\delta^2 / VAR(B_i)\,]$$

where σ_δ^2 is an estimate of the $\{B_i\}$ measurement noise. Since the sample data in Table 1 have zero covariances for all the appropriate ignored combinations and $\mu_A = \mu_B = 1$, we can check these results:

$$m_{AB} = 1 * [2/3 / (2/3 + 0.04)] \approx 0.943 \text{ and } \mu_A/\mu_B = 1/1 \approx 0.943 / [1 - 0.04/(2/3+0.04)]$$
$$m_{BA} = 1 * [2/3 / (2/3 + 0.16)] \approx 0.806 \text{ and } \mu_B/\mu_A = 1/1 \approx 0.806 / [1 - 0.16/(2/3+0.16)]$$

This would actually be a circular algebraic result if we did not know the measurement noise directly. One cannot use the data for A and B and an assumption that they are related to the truth with the same slope (and thus to each other with unit slope) to find estimates of the variances of the noise, and also use the same noise estimates to correct the slopes from linear regression fits. One either has external information (or assumes) that they are both related to the truth with the same (or at least nearly the same) slopes and then finds estimates of the variances and noise as in Subsection 1.1, or one has external estimates of the measurement noise and then uses them to estimate the truth variability and to correct the slope estimates.

This completes the explanation of the slopes of the fit lines in Figure 1, and provides a means of correcting or estimating their inadequacies. An already obvious result is that one should use the less noisy variable (noise size relative to its variance) as the independent variable if presenting a single standard fit line. In the absence of information on which of two match-up measurements are noisier, one could assume that they have equal-size relative noise contributions, and thus the slope estimates will be equally biased (in relative absolute value) toward zero, and use the geometric mean of one with the reciprocal of the other as an estimate of the true slope. That is, if $VAR(\varepsilon_i)/VAR(A_i) \approx VAR(\delta_i)/VAR(B_i)$, then

$$\mu_A/\mu_B \approx SQRT(m_{AB}/m_{BA}) \quad \text{or} \quad \mu_B/\mu_A \approx SQRT(m_{BA}/m_{AB}).$$

Since

$$SQRT(m_{AB}/m_{BA})=SQRT[VAR(A_i)/VAR(B_i)] \quad \text{and} \quad SQRT(m_{BA}/m_{AB})=SQRT[VAR(B_i)/VAR(A_i)],$$

this simple result just gives the ratios of the two measurements' standard deviations as the slope estimate. It is forced by the equal relative noise assumption.

Absent any information on which variable is noisier, one could also construct a bounding interval for the slope estimates by using the direct estimate and the reciprocal of the alternative fit estimate, that is,

$$\mu_A/\mu_B \in [m_{AB},1/m_{BA}] \quad \text{or} \quad \mu_B/\mu_A \in [m_{BA},1/m_{AB}].$$

One could be more conservative and extend these intervals by the standard uncertainties in the regression fit slope estimates.

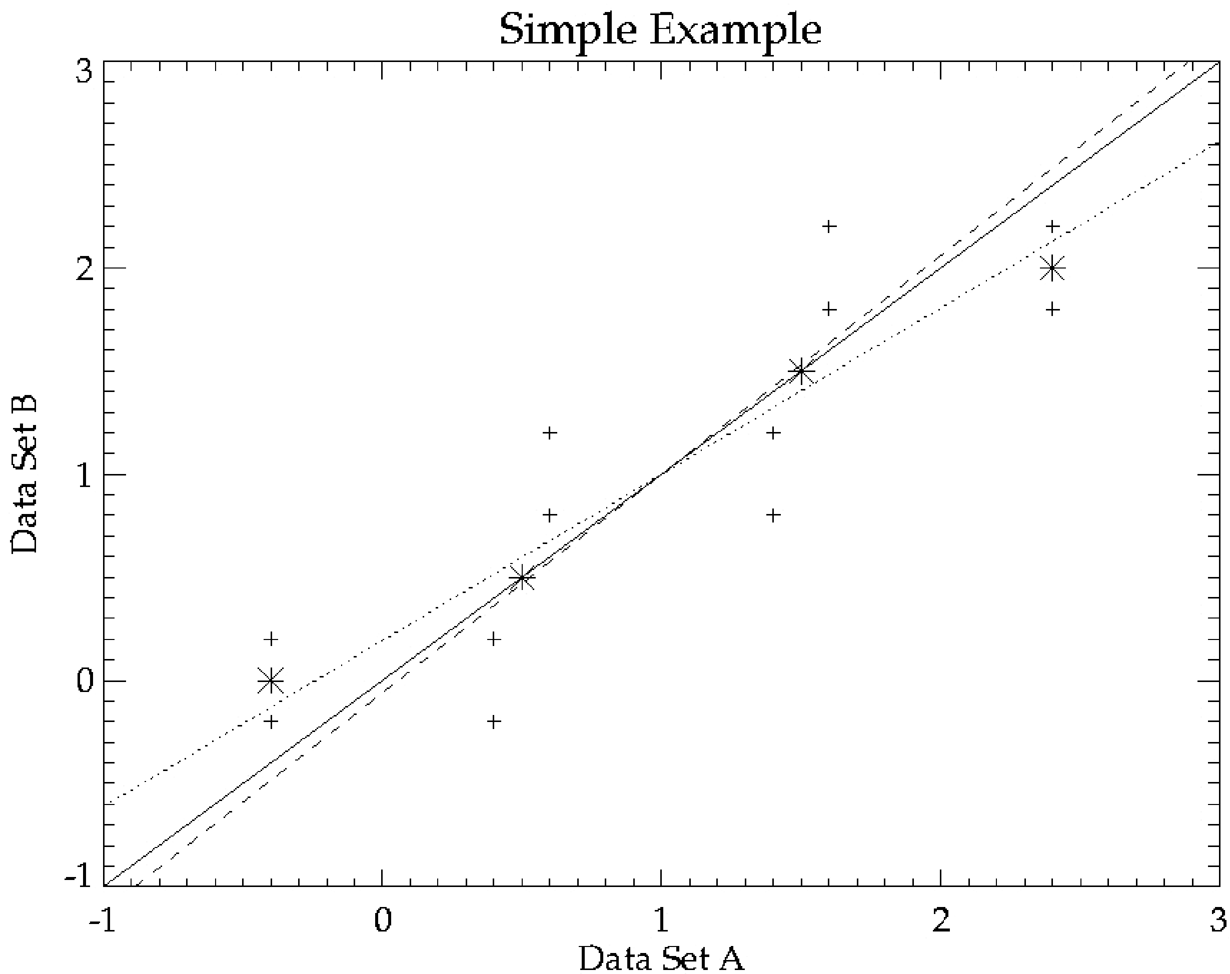
Finally, if we take the product of the estimates for m_{AB} and m_{BA}, we find

$$m_{AB} * m_{BA} = COV(A_i,B_i)/VAR(B_i) * COV(B_i,A_i)/VAR(A_i) = COV^2(A_i,B_i)/[VAR(B_i)*VAR(A_i)]$$
$$= \rho^2(A_i,B_i) \leq 1$$

where ρ is the Pearson linear correlation coefficient of $\{A_i\}$ and $\{B_i\}$. In the absence of noise in either variable this product should equal 1 since the two actual slopes are reciprocals. This is an alternative way to demonstrate that the slope estimates are biased toward smaller absolute values, and that the geometric mean of the relative accuracy for the two slope estimates is the Pearson linear correlation coefficient.

Subsection 1.3. Binning with Noisy Independent Variables

The ✳ symbols in Figure 1, reproduced below, plot the bin-averaged values obtained by finding the average of the $\{B_i\}$ data values in four bins, {[-1,0), [0,1), [1,2), and [2,3)}, using the $\{A_i\}$ data to determine the bins. Even though the B data were constructed to have absolutely no measurement bias (The average value for the B measurements of any truth value is the truth value.), it appears that B values are too high for small truth values and too low for large truth values. The source of this misleading result can be seen in the use of the A values to determine the bins. The A data in the lowest bin, [-1,0), get there because the A measurement noise is negative, while the mean of the B data in this bin is the correct truth average value of 0 (The B data errors are uncorrelated with the A errors.); similarly, the A data in the highest bin, [2,3), get there because the A measurement noise is positive. This type of false bias can be accentuated when the truth data distribution drops off in density at the extreme values, since the small number of truth values properly in an extreme bin can be overwhelmed by a larger number of measurements in the bin due to noise in the bin-selecting measurement.



Repeat of Figure 1.

When comparing two noisy measurement sets without truth values available to determine the binning, one should see how the results change with the binning variable switched and also with the binning performed using the averages of the paired values to sort the data. An alternative approach is to compare the distributions of measurements from the two systems – effectively discarding the match-up pairing information but keeping the overall similar sampling of annual or other cycles and real variations. This is considered in Section 3. Appendix 2 considers comparisons of histograms. Measurement noise can also cause difficulties for computing percent differences in a related manner as that seen with binning examples (in addition to the commonly observed problems from increased nonlinearity of percent calculations when the size of the noise rises to an appreciable fraction of the truth values' magnitude), so care should be taken in the choice of a reference variable there as well.

Subsection 1.4. Caveats

At times in this section, we have assumed that the two measurements are related to the truth with a slope of unity (or at least with identical slopes) and that noise in the chosen "independent" variable is responsible for biased results. Unfortunately, it is also common that a measurement system really does a less than ideal job of capturing the variability of the truth, that is, the measurements have reduced variance because of a lack of sensitivity and estimates of slopes from regression fits are correctly less than 1, rather than because of noise in the independent variable choice. Data from maximum-likelihood retrievals are a case in point where one may not expect to retrieve the full truth variance because the values are biased toward an *a priori*. One should consider the analysis tools, for example, averaging kernels and retrieval covariance versus *a priori* covariance (See, *e.g.*, Rodgers, 1990), available for such retrievals to try to get independent estimates of the measurement-to-truth relationship. The sensitivity of the different physical measurement to real changes should also be considered.

Subsection 1.5. Section Summary
In standard linear regression problems, one has a set of noisy measurements related to a set of non-noisy independent values. If the independent values are also noisy, then there is not enough information to uniquely estimate the linear relation between the two sets without additional assumptions. The Pearson linear correlation coefficient gives one measure of the range of slopes which are consistent with such measurements in the absence of other information on the noise. Any external piece of information giving the value of one of the following six quantities: 1. or 2. the signal variability captured in a measurement set, $\mu_A^2 VAR(R_i)$ or $\mu_B^2 VAR(R_i)$; 3. or 4. the noise variability present in a measurement set, σ_ε^2 or σ_δ^2; 5. the relative signal (slope), μ_A/μ_B; or 6. the relative noise variability, $\sigma_\varepsilon^2/\sigma_\delta^2$, allows one to determine estimates for the other five from the following quantities

$$VAR(A_i) = \mu_A^2\, VAR(R_i) + 2*\mu_A COV(R_i,\varepsilon_i) + VAR(\varepsilon_i) \approx \mu_A^2\, VAR(R_i) + \sigma_\varepsilon^2$$
$$VAR(B_i) = \mu_B^2\, VAR(R_i) + 2*\mu_B COV(R_i,\delta_i) + VAR(\delta_i) \approx \mu_B^2\, VAR(R_i) + \sigma_\delta^2$$
$$COV(A_i,B_i) = \mu_A\mu_B VAR(R_i) + \mu_B COV(R_i,\varepsilon_i) + \mu_A COV(R_i,\delta_i) + COV(\varepsilon_i,\delta_i) \approx \mu_A\mu_B VAR(R_i)$$
$$VAR(A_i-B_i) = VAR(A_i) - 2\, COV(A_i,B_i) + VAR(B_i) \approx (\mu_A-\mu_B)^2 VAR(R_i) + \sigma_\varepsilon^2 + \sigma_\delta^2$$

These results also rely on an assumption that the noise terms in the two measurements are uncorrelated with each other and with the truth variability.

SECTION 2. A More Interesting Example

In order to show some of the difficulties one can encounter in trying to apply the formulas in the previous section, we create some synthetic data sets using a random number generator. This gives us knowledge of the truth data and control of the noise statistics. (Appendix 3 has an IDL program using a random number generator to create synthetic, pseudo-normally-distributed noise values for the examples in this section.)

We create a uniformly-distributed non-random data set of 401 data values in the interval [3,5]. This is the truth data set, $\{R_i\}$. We then use a random number routine to create pairs of 401-element sets of pseudo-normally-distributed zero-mean noise one with a standard deviation of 1/3 and the other with a standard deviation of 1/4. We add these to the truth data to form the $\{A_i\}$ and $\{B_i\}$ measurement sets. By construction, the models have $\mu_A=\mu_B=1$, $\mu_B/\mu_A=\mu_A/\mu_B=1$, a=b=c=d=0, and $\sigma_\varepsilon^2=1/9$ and $\sigma_\delta^2=1/16$. Table 2.a and Figure 2 give some statistics and a scatter plot for a single realization of this experiment.

Table 2.a. Variance, Covariance, and Slope Computations.

	Computed-Variances				Est.-Var.		Estimated-Slopes			Pearson	Interval
	V_A	V_B	V_{A-B}	CV_{AB}	V_A–CV	V_B–CV	CV/V_A	CV/V_B	$\sqrt{(V_B/V_A)}$	CV/$\sqrt{(V_BV_A)}$	
					δ=A-R	ε=B-R	m1	m2		ρ	
Case									$\sqrt{(m1/m2)}$	$\sqrt{(m1*m2)}$	[m1,1/m2]
Model	0.446	0.398	0.174	0.335	0.1111	0.0625	0.75	0.84	0.94	0.796	[0.75,1.19]
Single*	0.443	0.421	0.178	0.343	0.099	0.078	0.78	0.82	0.98	0.795	[0.78,1.23]
Direct*					0.1114	0.0638					
Col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)

The standard errors for the slope estimates are approximately 0.03.
***Single Case with ISEED=100675**

Table 2.a gives information on the estimates for the slopes and related quantities. The first case gives the values expected from the model design and noise contributions; the second case gives the actual statistics for a single realization of the experiment. The last eight columns of values are computed from the first three columns' quantities; VAR(A_i), VAR(B_i), and VAR(A_i–B_i). Columns 4, 5 & 6 give the Subsection 1.1 estimates of $\mu_A\mu_B$VAR(R_i), VAR(δ_i), and VAR(ε_i), respectively, assuming that the relative slope is 1. Since we know $\{R_i\}$, the A–R and B–R variance values can be calculated directly. These are given in the row for the third case. For this particular realization, the variance of the $\{B_i\}$ set is larger than expected from the model, leading to errors in the estimates from Subsection 1.1 calculations. (See the red entries in Table 2.) The direct calculations show that this is primarily caused by random correlations in terms dropped in the approximations in Subsection 1.1, not a real increase in the noise variability for the $\{B_i\}$ measurements.

Given the size of the noise in the independent variables and Subsection 1.2 analysis, we expect standard linear fit slopes of 0.751 for the first case and 0.843 for the second (Columns 7 and 8). Alternatively, adjusting the fit slopes by using the prescribed model variances and the Subsection 1.2 corrections will produce consistent new slope estimates of 1.032 for the first case and 0.967 for the second; both are closer to the 1.000 model slopes than the unadjusted fit values. Notice that the line with the true slope, μ_B/μ_A=1.000, lies between the two fit lines. The geometric mean of the first fit slope times the reciprocal of the second fit slope (Column 9) is SQRT(0.775/0.815) ≈ 0.98. This is closer to the true slope than one would expect given the uneven model noise. One expects to find an equal-noise slope estimate of SQRT(0.751/0.843) ≈ 0.94 because of the actual difference in the sizes of the noise components between the two sets. The Pearson linear correlation coefficient (Column 10) for this sample is 0.795, and equals the geometric mean of the two slope estimates SQRT(0.775*0.815) = 0.795 . In the absence of any information on the noise in the measurements one could create an interval by using one of the slopes and the reciprocal of the other of [0.78,1.23] (Column 11). This interval could be extended even further considering the 0.03 uncertainty in the slope estimates from standard residual analysis.

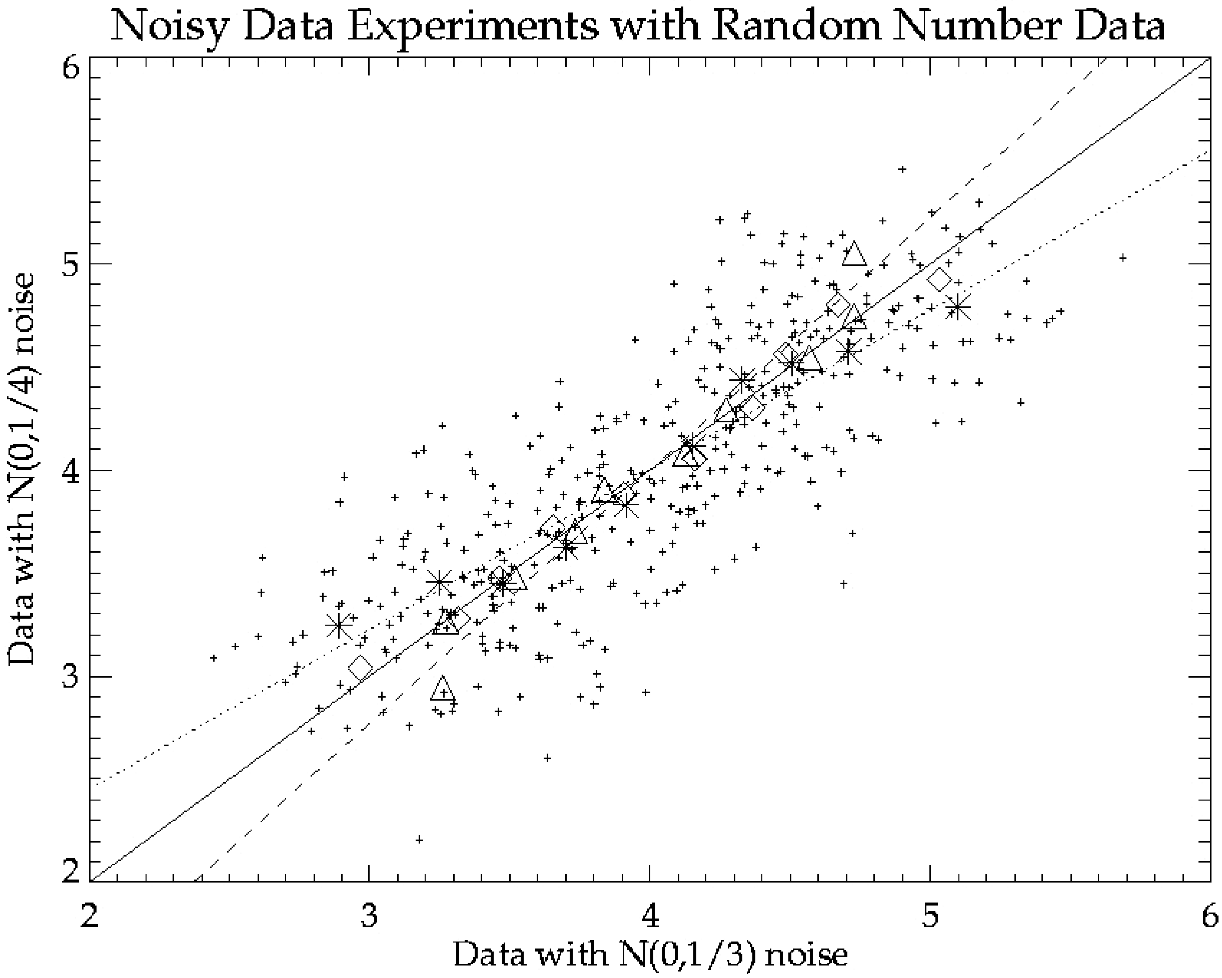


Figure 2. Sample results for the single case of synthetic data used in Table 2.a. The solid reference line is B=A. The A-axis is horizontal and the B-axis is vertical. The dotted line is from a linear regression fit of the B values with the A values. It has a B/A-slope of 0.775, a B-intercept of 0.900 and passes through the point (4.006,4.006). The dashed line is from a linear regression fit of the A values with the B values. It has an A/B-slope of 0.815 and an A-intercept of 0.742 and passes through the point (4.006,4.006). Figure 2 also shows the results of 10-percentile binning. The ◇ symbols display the average values for the data binned by using the sum of the A and B values to determine the 10-percentile bins, the Δ symbols display the average values for the data binned by using the B values, and the ✳ symbols display the average values for the data binned by using the A values. (Case: ISEED=100675)

The three lines on the scatter plot in Figure 2 are as in Figure 1. The solid reference line is just A=B. The dotted line is from a linear regression fit of the B values with the A values. It has a B/A slope of 0.775, a B-intercept of 0.900, and passes through the (A,B) point (4.000,4.001). The dashed line is from a linear regression fit of the A values with the B values. It has an A/B slope of 0.815, an A-intercept of 0.742, and passes through the (B,A) point (4.001,4.000).

Figure 2 also shows the results of 10-percentile binning. The ◇ symbols display the average values for the data binned by using the sum of the A and B values to determine the 10-percentile bins, the Δ symbols display the average values for the data binned by using the B values, and the ✳ symbols display the average values for the data binned by using the A values. Notice that the ◇ symbols are much closer to the B=A line for the extreme bins than the points for the other two binning alternatives. Taken alone, either the Δ or ✳ symbols would lead one to believe, incorrectly, that there was a problem with the ability of one of the measurement sets to estimate extreme values. Consider the dramatic difference in the presentation if only the dotted fit line and the ✳ symbols were shown on Figure 2.

The simulation code was used to generate a Monte-Carlo experiment with 5000 repetitions of this simple experiment. The first case in Table 2.b gives the values expected from the model design and noise contributions, the second case uses the estimates from the 5000 repetitions from Subsection 1.1 results, with an assumption that $\mu_C=\mu_A/\mu_B=1$. The 5000 repetitions gave Subsection 1.1 overall average estimates of the truth and noise variances with little bias (0.3357, 0.1112, and 0.06249 versus the corresponding model parameter values of 0.3350, 0.1111, and 0.0625) and reasonable standard deviations (0.0128, 0.0150, and 0.01327) for the truth and A and B noise components, respectively. The particular case presented in Table 2.a was typical of the Monte Carlo results in the magnitude of its differences from the model predictions.

Table 2.b. Section 1.1 Estimates from Monte Carlo Study.

	R	δ=A-R	ε=B-R
Model Variance	0.3350	0.1111	0.06250
Avg. Variance[MC]	0.3357	0.1112	0.06249
S.D. of Variance[MC]	0.0128	0.0150	0.0133

[MC]Results for 5000 Monte Carlo repetitions

Even though the data in this section are less contrived than that in the previous section, it is still an ideal case because of the perfect match-up between the two sets of measurements and because the two sets of measurements completely capture the truth variability before the addition of independent measurement noise. In general, there will be differences between the match-up measurements because they are not measuring at the same place or time. This will create variability between the two measurements. It is optimistic to think that this variability may be simply included in the noise. Given a match-up between a ground-based measurement and a satellite-based measurement, for example, one could argue that the satellite should see larger truth variability because it samples in many different places, or one could argue that it should see less variability because its measurements average the conditions over larger regions. The method in Subsection 1.1 will interpret increased variability in the underlying truth data sampled by one system relative to another as measurement noise in the more variable truth system (assuming that the increased variability is not correlated with the truth variability). As we will see again in the next section, the assumption that both measurements have the same slope versus the truth (the same sensitivity to variations in the truth) is a key condition.

SECTION 3. Practical Examples of Match-up Data Sets

This section presents some results for a real-world example. The data come from comparisons between total column atmospheric ozone measurements derived from ground-based and satellite-based instruments. We use an overpass data set of retrieved ozone estimates from the NOAA-16 SBUV/2 (Solar Backscatter UltraViolet instrument, http://www2.ncdc.noaa.gov/docs/klm/html/c3/sec3-8.htm) measurements for the last five years compared to the estimates from the Mauna Loa Dobson station measurements. (http://www.mlo.noaa.gov/programs/esrl/dobsonws/dobsonws.html) There are a total of 968 days where measurements were taken by both systems, satisfying the match-up criteria. (The SBUV/2 data values are inverse-distance-weighted averages of all Version 8 retrievals within 2° of latitude and 20° of longitude of the Mauna Loa Observatory, made within 12 hours of the Dobson measurements.)

There are actually two different total column ozone estimates computed by the Version 8 SBUV/2 retrieval algorithm. One, best total ozone (http://toms.gsfc.nasa.gov/version8/v8toms_atbd.pdf), uses ratios of measurements at pairs of wavelengths and a look-up table with a range of standard profiles and climatological adjustments. The other, profile total ozone, creates ozone profile estimates by using a range (See ftp://www.orbit.nesdis.noaa.gov/pub/smcd/spb/ozone/dvd_v8/DVDhtml/V8_Algorithm_Description.html.) of ozone sensitive wavelengths and seasonally varying *a priori* ozone profiles in a maximum-likelihood retrieval. The match-up SBUV/2 data were screened against each other to identify unrealistic outliers, and one value was removed from each set. The Mauna Loa Dobson station is located far above the surrounding Pacific Ocean (3.4 KM altitude). The large field-of-view (approximately 200 KM by 200 KM) of the SBUV/2 and the low likelihood of a perfect match-up mean that the satellite measurement usually includes a column to sea level contribution not seen by the Dobson instrument. To lessen the impact of this effect, the satellite data were initially adjusted by subtracting 12 DU from each value before comparisons with the Dobson data.

Table 3 gives some statistics on four different match-ups between the three data sets. Scatter plots comparing the match-up data sets appear in Figure 3. Each scatter plot also has a reference line with a slope of 1 and the two linear regression fit lines.

Table 3. Statistics, Variances and Slopes for Match-Up Data.

Key: G≡Ground-based, Sp≡SBUV/2 Profile Total, Sb≡SBUV/2 Best Total, Spa≡Adjusted Sp

Mean 257.9 262.6 260.1 259.2

Case	Computed Variances				Est. Var.		Estimated Slopes			Pearson	Interval
	V_A	V_B	V_{A-B}	CV_{AB}	V_A-CV	V_B-CV	CV/V_A	CV/V_B	√(V_B/V_A)	CV/√(V_BV_A)	
					δ	ε	m1	m2		ρ	
C1	Sp	G	Sp-G	COV	Sp-R	G-R	G/Sp	Sp/G	√(m1/m2)	√(m1*m2)	[m1,1/m2]
	351.7	331.0	57.9	312.4	39.3	18.6	0.89	0.94	0.97	0.916	[0.89,1.06]
C2	Sb	G	Sb-G	COV	Sb-R	G-R	G/Sb	Sb/G			
	304.6	331.0	63.4	286.1	18.5	44.9	0.94	0.86	1.04	0.901	[0.94,1.16]
C3	Spa	G	Spa-G	COV	Spa-R	G-R	G/Spa	Spa/G			
	331.7	331.0	50.2	306.3	25.5	24.7	0.92	0.93	1.00	0.924	[0.92,1.08]
C4	Sb	Sp	Sb-Sp	COV	Sb-R	Sp-R	Sp/Sb	Sb/Sp			
	304.6	351.7	10.5	322.9	-18.3	28.8	1.06	0.92	1.08	0.987	[1.06,1.09]
Col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)

Means are in DU with Sp, Sb and Spa adjusted as described in the text.
Variances and covariances are in DU^2. Slopes are non-dimensional – in DU/DU.
The standard errors for the slope estimates are approximately 0.013 for the cases involving G and approximately 0.005 for the fourth case.

The four match-up cases are as follows: the first case, C1, is for the SBUV/2 profile total (Sp) versus the Dobson total (G); the second, C2, is for the SBUV/2 best total (Sb) versus Dobson total (G); the third, C3, is for the SBUV/2 profile total with an additional adjustment (Spa) versus the Dobson total (G); and the fourth, C4, is for the SBUV/2 best total (Sb) versus profile total (Sp). The first four columns in Table 3 give the computed values for the variances used in Subsection 1.1, namely, the variances of both of the sets and their differences and their covariances. Notice that the SBUV/2 best total (Sb) is much less variable than the SBUV/2 profile total (Sp) and that the Dobson total (G) result is in between. Also notice that the two SBUV/2 sets have very little variance in their differences (fourth case, third column in red). The covariances (fourth column, first three cases) of the ground-based measurement with the satellite-based ones range from 286.1 to 312.4. The next two columns present the estimates of the noise variances computed by using the Subsection 1.1 formulas with $\mu_A/\mu_B = 1$, that is, assuming that the slopes versus the truth are equal. Notice that the Dobson noise estimates (sixth column, first three cases) range from 18.6 to 44.9. The fourth case gives a non-physical result: a negative variance for the SBUV/2 Sb noise of **-18.3**.

The seventh and eighth columns give the slope results of standard regression fits of one set to the other. The seventh uses the first column variable as the independent variable and the eighth uses the second column variable as the independent variable. The ninth column gives the geometric mean of the first slope and the reciprocal of the second slope (equal to the square root of the ratio of the second variance to the first), and the tenth column gives the Pearson linear correlation coefficient (equal to the square root of the product of the slope estimates). The eleventh column gives the slope interval using the first slope estimate and the reciprocal of the second one as the endpoints. The standard errors are given in the note below the table. Notice for the fourth case that the first slope is greater than 1, and that the slope interval (Column 11) does not contain 1.

Figures 3, 4, & 5 appear here and on the next two pages each, with four plots corresponding to the four cases in Table 3 with C1 top/left, C2 top/right, C3 bottom/left, and C4 bottom/right. Differences in the results for the three binning options are displayed in Figure 4, and differences using the sorted data distributions are in Figure 5.

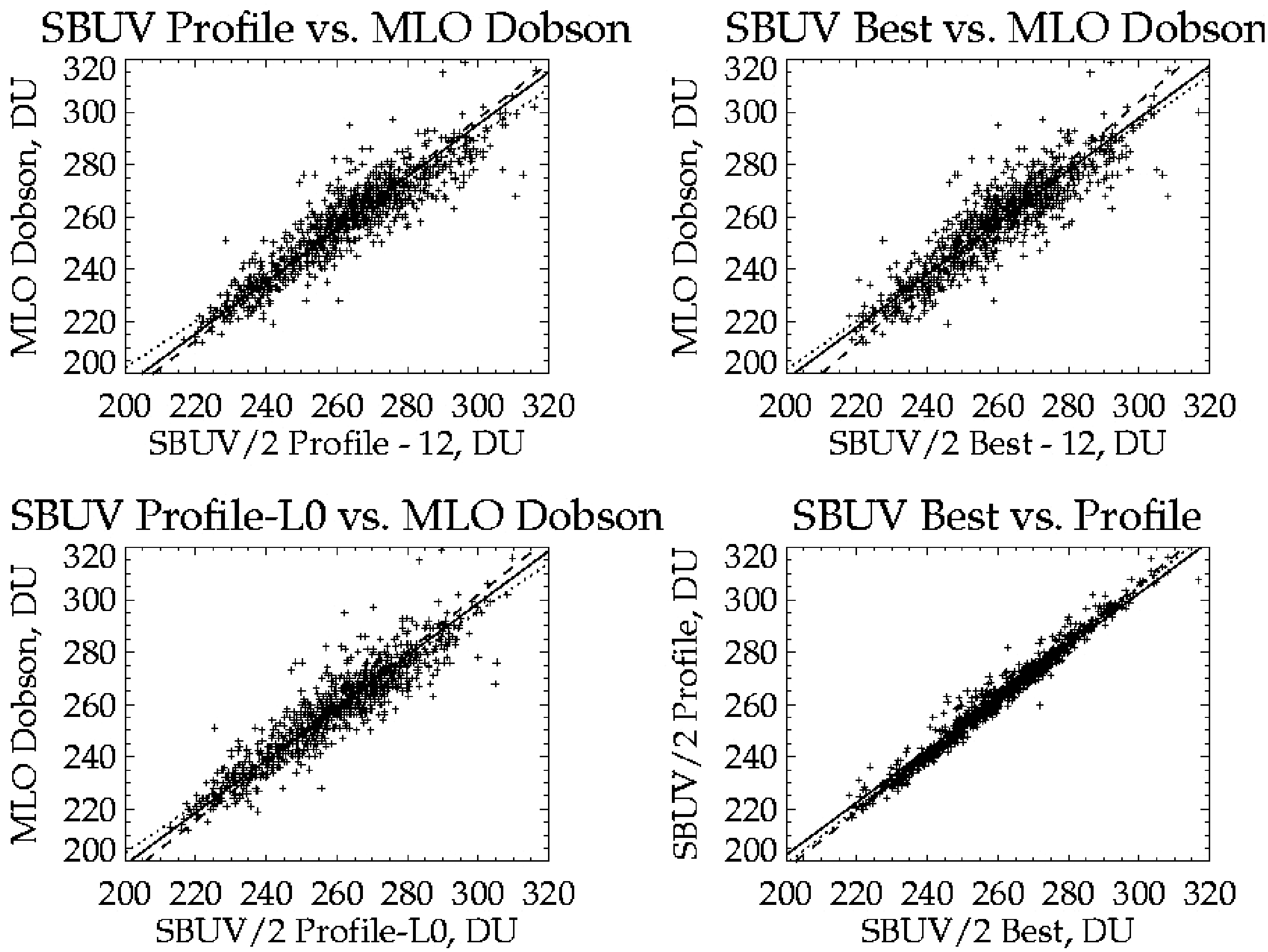


Figure 3. Scatter plots for the four cases along with the two standard fit lines – the dotted lines use the horizontal data as the independent variable, and the dashed lines use the vertical data as the independent variable. Slopes are given in the eighth and ninth columns in Table 3. The solid line is a line with slope 1 passing through the centroid of the data; (average of horizontal,average of vertical). The Pearson linear correlation coefficients are: 0.916 for Sp with G, 0.901 for Sb with G, 0.924 for Spa with G, and 0.987 for Sb with Sp. One can show algebraically that this correlation coefficient is equivalent to the square root of the product of two least-squares-fit slope estimates. Notice on the plot in the lower right, that both fit lines go from below the solid reference line for low values to above it for high values, unlike the other cases. For consistency of presentation, the Sp and Sb values in the fourth case have had 12 DU subtracted from each value as was done for the comparisons in the first and second cases.

Figure 4 shows the differences between binned-averages for each pair of vertical and horizontal values from Figure 3. The eleven points are each averages of 88 binned values for the Figure 3 data binned by using the vertical (Δ), the horizontal (✳), or the sum of both (◇) values as the binning determinant. Figure 5 shows the differences of the distributions for each pair of data in Figure 3. The data sets are sorted independently from smallest to largest and the difference in the corresponding ordered values are plotted. The overall mean differences are adjusted to zero in Figures 4 & 5.

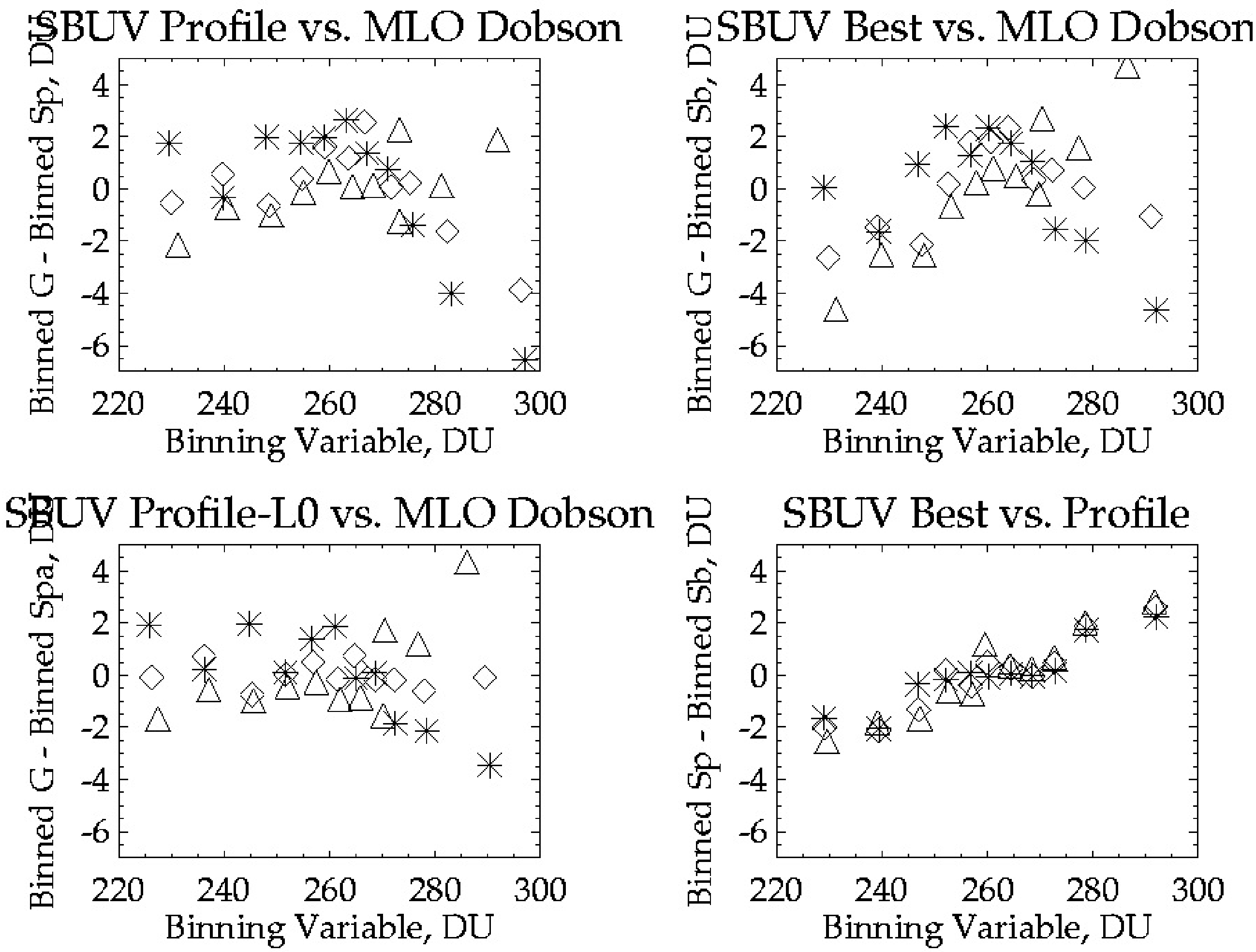


Figure 4. Differences of binned data values for the four match-up cases in Figure 3. The eleven points are each averages of 88 binned values for the Figure 3 data. The data set match-up pairs are sorted into these bins by using the vertical (Δ), the horizontal (✳), or the sum of both (◇) values as the binning determinant. Regardless of how the points are sorted into bins to create the bin-averages, the differences for the vertical axis variable are computed with the same choices for which one is the minuend and which one is the subtrahend.

We begin with a discussion of the SBUV/2 best ozone (Sb) versus SBUV/2 profile total (Sp). This case is the last one in Table 3 and the plots appear in the lower right panels of Figures 3, 4, & 5. The data for these two sets are from very good match-ups (in terms of location and time) as they are usually for retrievals of the same set of radiance measurements. (The SBUV/2 makes a sequence of measurements at 12 wavelengths over 24 seconds. The profile retrievals for this latitude come from information in the measurements at the shortest eight wavelengths and the best total comes from information in the longest four wavelengths.) While the difference in the variances is large (304.6 DU^2 versus 351.7 DU^2) the variance of the differences is small (10.5 DU^2). This means that the standard deviation of the match-up differences is less than 3.3 DU. (Remember the variance of the differences is an upper bound on the sum of the noise variances of the two measurements.) It also means that much of the additional variability in the Sp data must be correlated with the variability in the Sb data.

The Subsection 1.1 equal-slope estimate of the truth variance is greater than the best variance, leading to an impossible negative value for the estimate of the Sb noise variance. The slope of the fit using the best ozone (Sb) as the independent variable is greater than 1 (specifically 1.060) and the reciprocal of the slope for the profile ozone (Sp) as the independent variable is 1.089. We expect the first one to be less than the true slope and the second one to be greater than the true slope because of noise in the independent variables. So the problem is that the two sets have different slopes compared to the truth, that is, that μ_A is not equal to μ_B as assumed in Subsection 1.1 or, equivalently, that μ_A/μ_B is not equal to 1 as assumed in Subsection 1.2, but that μ_A/μ_B probably lies somewhere in the interval [1.060,1.089]. This means that the profile ozone total has a larger slope relative to the truth than the best total ozone. (It does not tell us which one has a slope closer to unity relative to the truth.) This is confirmed in the plots in the lower right panels of Figures 4 & 5, where the results for binning by all three choices and the differences of the sorted data show the Sp with lower lows and higher highs than Sb. By examining the retrieval algorithms, one finds that the best ozone has more sensitivity to real variations in tropospheric ozone, while the profile ozone uses a set of seasonally-varying climatological *a priori* profiles that have considerable variability in the lower two layers, and that this imposed variability is correlated with the real total ozone variability. For this particular region, one can estimate a profile total result from the best total result by multiplying it by 1.075 and subtracting 17 DU, and this approximation is very good.

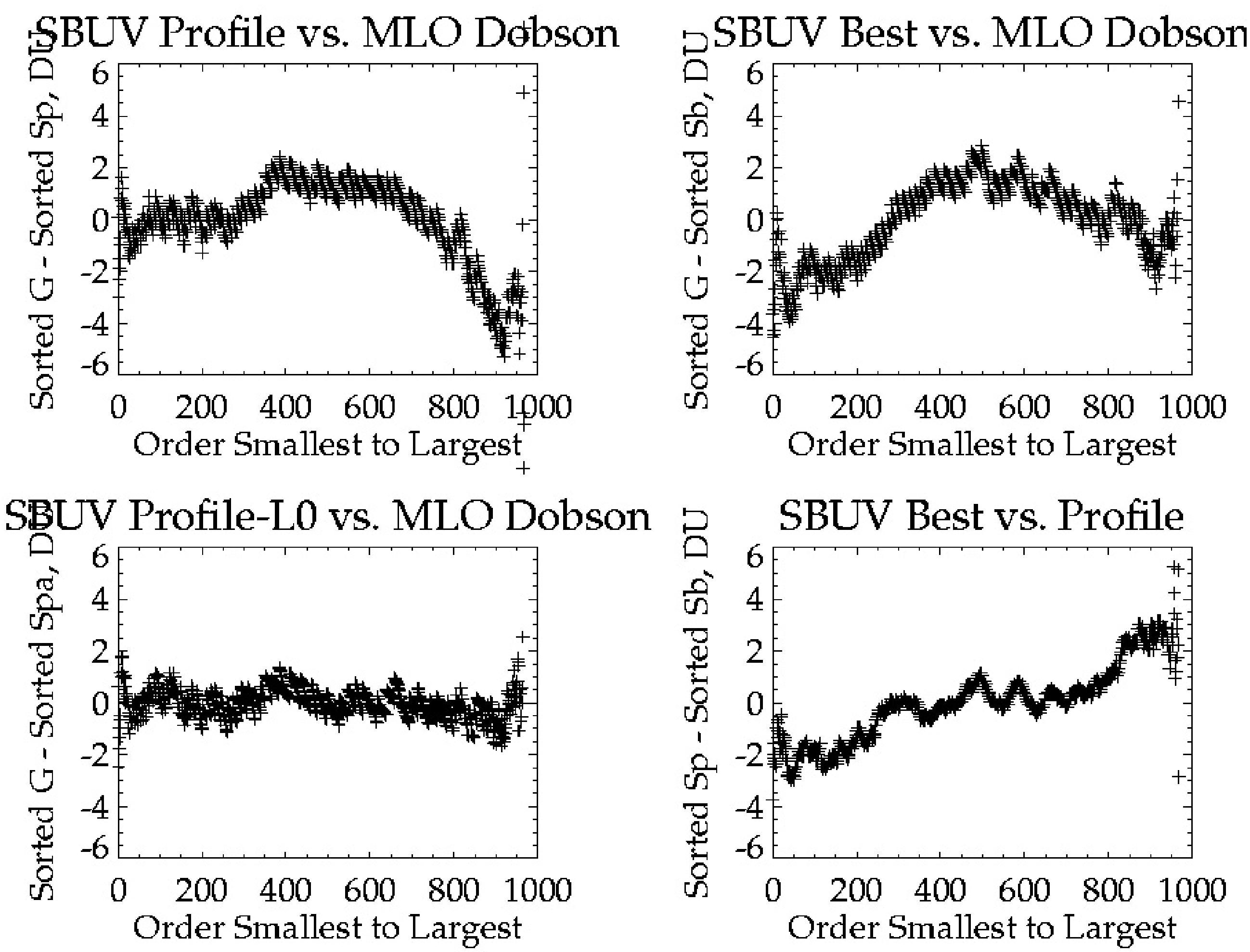


Figure 5. Differences of sorted data values for the four match-up cases in Figure 3. The data sets are sorted independently from smallest to largest and the differences (vertical values minus horizontal values) in the corresponding ordered values are plotted. The overall mean differences are adjusted to zero, that is, the biases between sets are removed. Each plot has 968 points. The patterned appearance and spread in the top two plots is related to the Dobson station discrete reporting precision of 1 DU.

Repeat of Table 3. Statistics for Match-Up Data.

Case	Computed Variances				Est. Var.		Estimated Slopes			Pearson	Interval
	V_A	V_B	V_{A-B}	CV_{AB}	V_A-CV	V_B-CV	CV/V_A	CV/V_B	√(V_B/V_A)	CV/√(V_BV_A)	
					δ	ε	m1	m2		ρ	
C1	Sp	G	Sp-G	COV	Sp-R	G-R	G/Sp	Sp/G	√(m1/m2)	√(m1*m2)	[m1,1/m2]
	351.7	331.0	57.9	312.4	39.3	18.6	0.888	0.94	0.970	0.916	[0.89,1.06]
C2	Sb	G	Sb-G	COV	Sb-R	G-R	G/Sb	Sb/G			
	304.6	331.0	63.4	286.1	18.5	44.9	0.939	0.86	1.042	0.901	[0.94,1.16]
C3	Spa	G	Spa-G	COV	Spa-R	G-R	G/Spa	Spa/G			
	331.7	331.0	50.2	306.3	25.5	24.7	0.923	0.93	0.999	0.924	[0.92,1.08]
C4	Sb	Sp	Sb-Sp	COV	Sb-R	Sp-R	Sp/Sb	Sb/Sp			
	304.6	351.7	10.5	322.9	-18.3	28.8	1.060	0.92	1.075	0.987	[1.06,1.09]

Variances are in DU^2 and Slopes are non-dimensional.

We now consider the results for the first and second cases. Taken individually without the fourth case information, one might make the following conjectures:
1. For the first case, C1: either the slope of the Dobson total to the profile total is in the interval [0.888,1.059], or the SBUV/2 profile total (Sp) is *more noisy* than the Dobson station total data and the slope is 1.00, or the SBUV/2 profile total and Dobson total are equally noisy and the slope is approximately 0.970 (< *1*). The ◇s in the plot for this case (upper left of Figure 4) show that the Dobson values are smaller than the profile total values for high totals, that is, the profile totals have a larger range. This is also observed in the plot of the differences in the distributions for this case (upper left of Figure 5).
2. For the second case, C2: either the slope of the Dobson total to the best total is in the interval [0.939,1.157], or the SBUV/2 best total (Sb) is *less noisy* than the Dobson total data and the slope is 1.00, or the SBUV/2 best total and Dobson total are equally noisy and the slope is approximately 1.042 (> *1*). The ◇s in the plot for this case (upper right of Figure 4) show that Dobson values are smaller than the best total values for low totals, that is, the best totals have a smaller range. This is also observed in the plot of the differences in the distributions for this case (upper right of Figure 5).

Notice that the equal-noise assumptions would lead to slopes for these two cases differing by an amount consistent with the fourth case's slope results (1.042/0.970 ≈ 1.075). While this is required algebraically, the small variance of Sb–Sp gives confidence that the slope estimates for the fourth case are accurate and leads one to expect that the noise estimates for Sp–R, Sb–R and G–R for the first and second cases should be more consistent. Remember, the 1.06 slope value is expected to underestimate (be a lower bound on) the true slope of Sp fit with Sb if Sb has any measurement noise at all.

The SBUV/2 profile totals (Sp) were adjusted by subtracting 12 DU from them to account for the ozone below the altitude of the Mauna Loa Observatory. The Version 8 ozone profile retrieval algorithm uses a seasonally (and latitudinally) varying *a priori*, which does not have a constant amount in the lowest layer (1013 hPa to 507 hPa, called Layer 0) for Mauna Loa's latitude. The optimal estimation profile retrieval has little measurement information in this layer and relies mainly upon the *a priori* values there. It turns out, for this location, that the fixed seasonal *a priori* ozone variations in this lowest layer are correlated with the total ozone amount. Adjusting the SBUV/2 profile totals (Sp) by subtracting the varying layer 0 *a priori* amounts from them, instead of a constant 12 DU, gives a set of values (Spa) which should be more representative of the column measured by the Dobson station. The third case presents the results for this adjusted set versus the Dobson values. The conjectures for these results are:
(3) For the third case, C3: either the slope of the Dobson total to the profile total is in the interval [0.923,1.081], or if the SBUV/2 adjusted profile total and Dobson total are equally noisy then the slope is 0.999 (~1). The ◇ symbols in the plot for this case (lower left plot of Figure 4) show a fairly flat set of

differences as do the distribution comparisons (lower/left plot of Figure 5). Notice the relationship between the ratio of the variance of the differences to the truth variance (From Table 3, Case 3, 50.2/306.3≈0.16) and the width of the slope interval, 1.081–0.923≈0.16. When comparing two measurement systems, there may be a trade-off between using locations with low match-up or measurement noise versus those with large variability in the truth values. The Pearson linear correlation coefficient can be used to quantify this trade-off in constraints on the slope estimates.

In summary: for the first case, Sb versus G, it is likely that the Dobson values are capturing more real variability than the SBUV/2 best total values; for the second and third cases, Sp(a) versus G, after proper adjustment for the lower tropospheric column, the two systems are providing very consistent estimates; and for the fourth case, the profile retrieval *a priori* is introducing additional variability in the lower layers into the retrievals that is not present in the best total ozone standard profiles and the profile retrievals sense real variations in this region with lower efficiency than stratospheric ozone variations. The SBUV/2 data was not screened for clouds which further obscure tropospheric ozone variations from the satellite measurements.

SECTION 4. Practical Examples of Match-up Data Sets Continued: Temporal Aggregation

For time series of match-up data, we may have another option – temporal averaging. For example, if there is a seasonal variability of the signal, then we may be able to greatly reduce the noise component of the measurements by creating monthly or other time-interval averages of the data without causing the same reduction in the measurement variance contribution from real changes. Recall that the errors in the slope estimates were of the form:

$$m_{AB} = COV(A_i,B_i)/VAR(B_i) \approx \mu_A/\mu_B\,[1 - VAR(\delta_i)/VAR(B_i)]$$

If we can average the data in such a way that the $VAR(\delta_i)$ is reduced by a larger fraction than the $VAR(B_i)$, then we will obtain a less biased estimate of the slope.

The data for the four cases were grouped by taking eight sequential values at a time and the resulting 121 average match-up points were re-analyzed as in Section 3. Figure 6 shows the time series for the four data sets. The small dots are the individual values and the ∗ symbols are the aggregated values. If the noise and match-up error terms are independent and are not serially correlated, then one expects the variance from these sources to be reduced to 1/8 the original values. While the averaged values have less range than the individual values, the reduction is a small fraction of the original range.

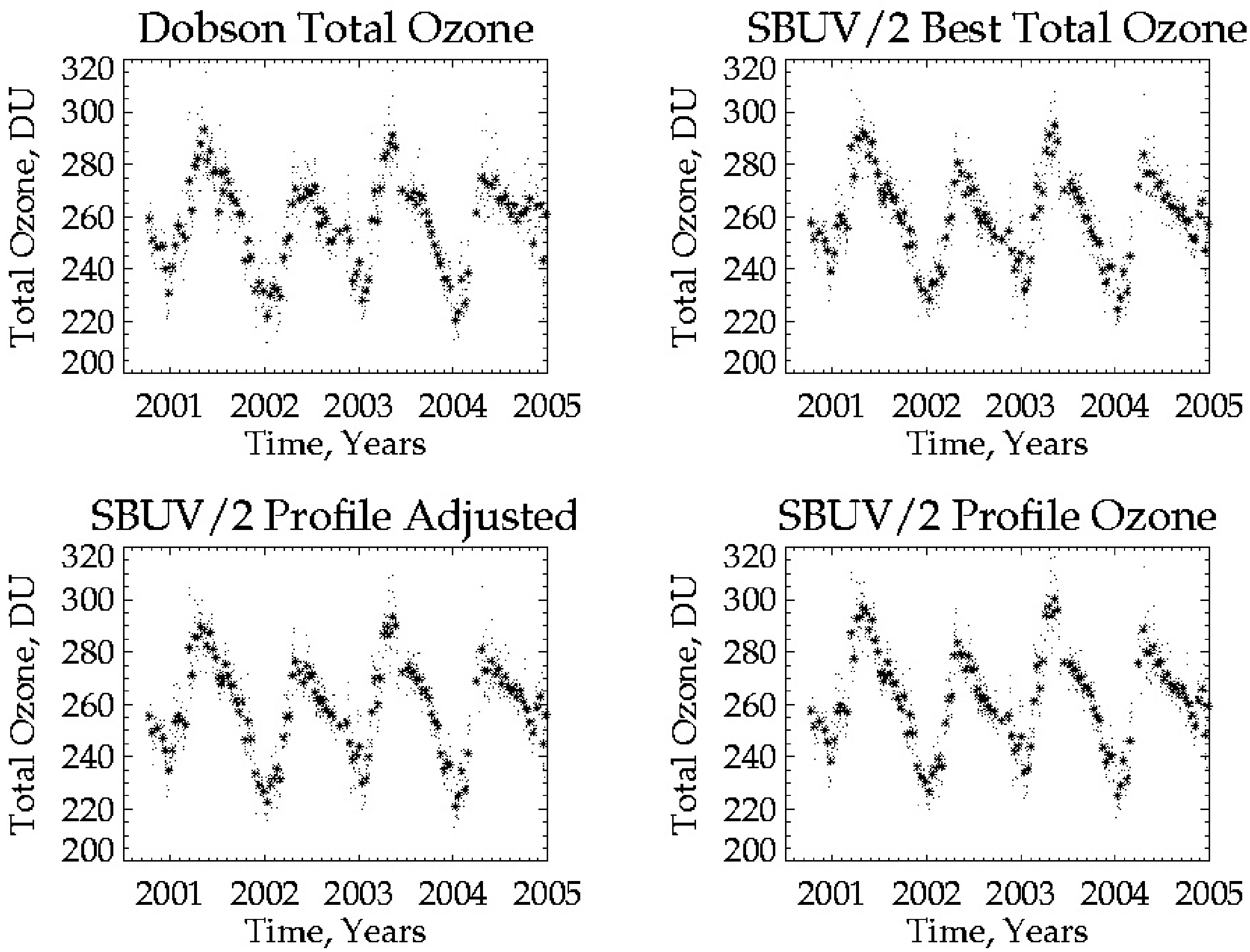


Figure 6. Plot of match-up time series with and without 8-point temporal aggregation. Small dots are 968 individual values used in the Section 3 analysis. The ✳ symbols are the 121 values for 8-point sequential averages.

The aggregated data were used to calculate the quantities in Table 3, and the results are given in Table 4 in the rows labeled "D8". The variances for the two sets of match-up data for each case, VAR(A_i) and VAR(B_i) (values in the first two rows), for the D8 cases are smaller than those for the non-aggregated, but are always over 0.8 times the original ones. On the other hand, recall that

$$\mathrm{VAR}(A_i - B_i) \approx (\mu_A - \mu_B)^2 \mathrm{VAR}(R_i) + \sigma_\varepsilon^2 + \sigma_\delta^2$$

so VAR(A_i–B_i) is an approximate upper bound on the total noise and match-up variance. Row three shows that this quantity is less than 0.4 of the original for the first three cases, and 0.6 for the fourth case. For the fourth case, there should be a significant contribution from the $(\mu_A - \mu_B)^2$VAR(R_i) term of approximately 2.5 DU^2. This means that the noise variance reduction is really from 8 DU^2 to 3 DU^2 which is similar to the other cases. Even for the third case where the two slopes are close in value, the reduction is not to 1/8 of the original covariance. One possible cause for this could be that there is a temporal drift in the calibration of one of the measurement systems which does not average out in the temporal aggregation. Even a 1% slowly-time-varying error would be large enough to produce these smaller reductions in the covariances ratios relative to the original ones. Another possible cause would be the imprecision of the linear model in representing the actual retrieval properties.

Table 4. Statistics for Match-Up Data and Temporal Averages of 8 Values.

Case	Computed Variances				Est. Var.		Estimated Slopes			Pearson	Interval
	V_A	V_B	V_{A-B}	CV_{AB}	V_A-CV	V_B-CV	CV/V_A	CV/V_B	$\sqrt{(V_B/V_A)}$	CV/$\sqrt{(V_B V_A)}$	
					Δ	ε	m1	m2		ρ	
C1	Sp	G	Sp-G	COV	Sp-R	G-R	G/Sp	Sp/G	$\sqrt{(m1/m2)}$	$\sqrt{(m1*m2)}$	[m1,1/m2]
	351.7	331.0	57.9	312.4	39.3	18.6	0.89	0.94	0.97	0.916	[0.89,1.06]
D8	313.6	279.2	21.3	285.8	27.8	-6.6	0.91	1.02	0.94	0.966	[0.91,0.98]
C2	Sb	G	Sb-G	COV	Sb-R	G-R	G/Sb	Sb/G			
	304.6	331.0	63.4	286.1	18.5	44.9	0.94	0.86	1.04	0.901	[0.94,1.16]
D8	262.9	279.2	21.8	260.1	2.7	19.1	0.99	0.93	1.03	0.960	[0.99,1.08]
C3	Spa	G	Spa-G	COV	Spa-R	G-R	G/Spa	Spa/G			
	331.7	331.0	50.2	306.3	25.5	24.7	0.92	0.93	1.00	0.924	[0.92,1.08]
D8	293.7	279.2	13.4	279.8	13.9	-0.5	0.95	1.00	0.98	0.977	[0.95,1.00]
C4	Sb	Sp	Sb-Sp	COV	Sb-R	Sp-R	Sp/Sb	Sb/Sp			
	304.6	351.7	10.5	322.9	-18.3	28.8	1.06	0.92	1.08	0.987	[1.06,1.09]
D8	262.9	313.6	5.5	285.5	-22.6	28.1	1.09	0.91	1.09	0.994	[1.09,1.10]
Col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)

Key: G Ground-based, Sp SBUV/2 Profile Total, Sb SBUV/2 Best Total, Spa Adjusted Sp
Variances are in DU^2 and Slopes are non-dimensional – in DU/DU.
The standard errors for the slope estimates for 8-point-averaged data are approximately 0.023 for the cases involving G and approximately 0.010 for the fourth case.

The Pearson correlation coefficients in the tenth column are closer to 1 for all four cases, and the slope intervals in the eleventh column are narrower even allowing for increases in the standard errors on the slopes. The new slope intervals all strengthen the arguments for the four cases in the previous section, namely, an ordering of Sp, G and Sb for response to signal variability, improved agreement for Spa with G, and a very accurate estimate of the linear relationship between Sp and Sb.

The relative decrease in the Dobson station data (G) variance from before to after aggregation was greater than for the other three sets. This could either be because it was noisier to begin with or because there were larger time-dependent changes (more drift) in the satellite data. Figure 7 shows the time series of differences between a data set and the fit using the slope in the ninth column for the four cases. The fourth case (lower right) shows a drift between the two SBUV/2 data sets of approximately 2 DU over the four years. There is also an increased variability in the middle of the record during a period with a known SBUV/2 instrument noise anomaly in Gain Range 3.

In addition to revealing (and responding to) possible time dependent drifts, the data aggregation may also change the emphasis of a comparison from short-term to long-term. That is, a measurement system may have different slopes or responses to (capabilities to resolve) short-term variability (*e.g.*, day-to-day) compared to longer term changes (*e.g.*, seasonal). For the current examples, one could fit each sets seasonal cycle with a functional form and use them to compare the original data indirectly. If the noise effects for the two sets on the seasonal cycle estimates were similar, then comparing these terms will avoid some of our current difficulties with an emphasis on the long-term variability reproduction. Additional time series and component analysis is beyond the scope of this report.

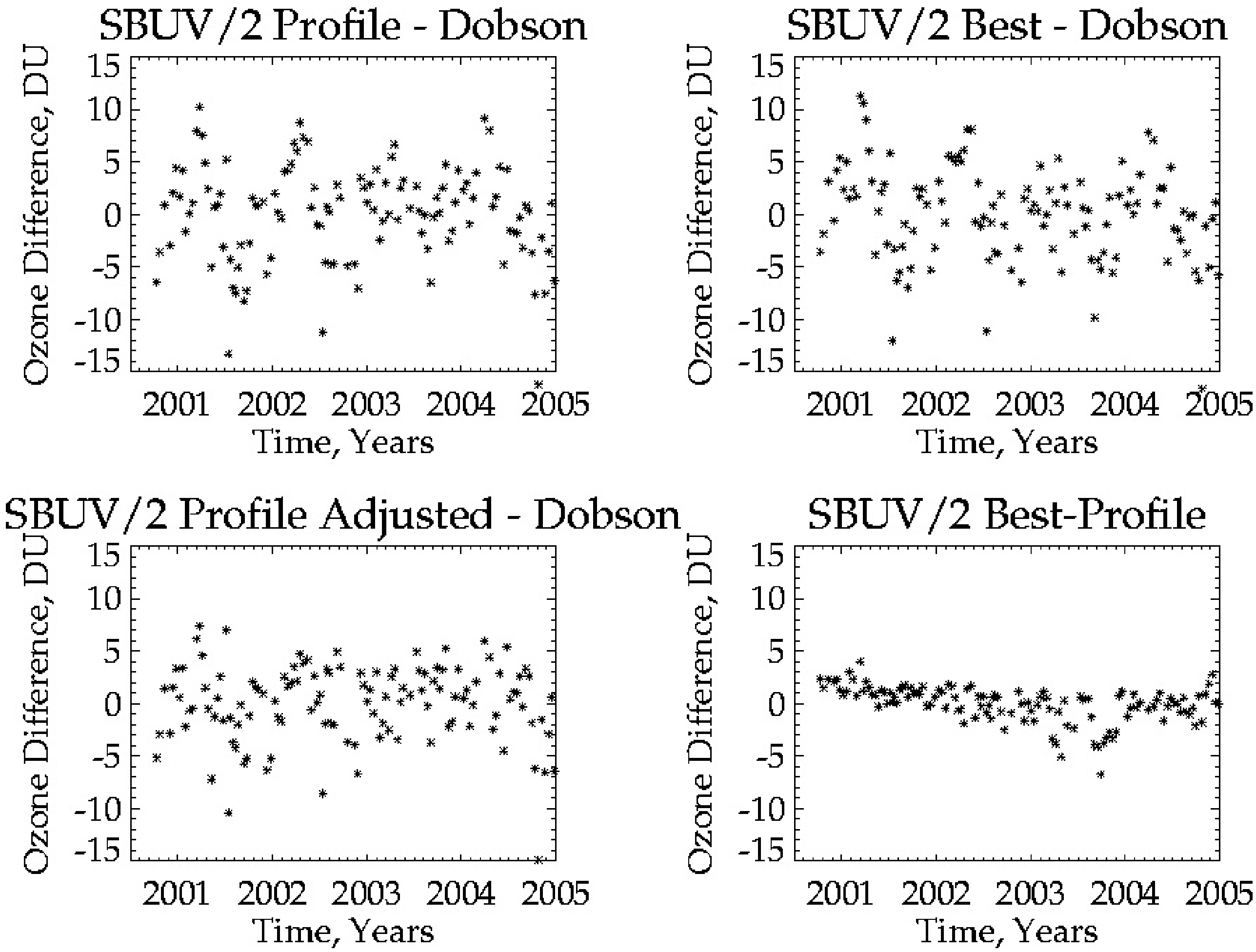


Figure 7. Time series of differences between 8-point aggregated data sets using linear fits. For each set, the difference is between the first value and a linear fit using the second value with the slope from an equal noise assumption for the independent and dependent variables, that is, slopes equal to the reciprocal of the estimates in Column 9 of Table 4 for the appropriate D8 row.

SECTION 5. Summary and Conclusions

This report has presented some simple analysis tools both to explain the failure of some common approaches used in the presentation of comparisons between noisy match-up data sets and to provide improved information on the true underlying behavior. A series of increasingly realistic examples has been explored to demonstrate some of the difficulties one can encounter even when well prepared. The key results can be summarized as follows: 1. The slope from a linear least squares regression fit of a dependent variable using a noisy independent variable will tend to be smaller in absolute value than the true slope between the two real non-noisy signals, and the magnitude of this decrease can be estimated given knowledge of the size of the noise in the independent variable; 2. The true linear relationship between a pair of noisy measurement data sets probably lies between the two standard fit lines (with each set taking on the independent variable role for one of the fit estimates); 3. Binning noisy data pairs by sorting based on the sums of the two match-up values may lessen false bias results in the extreme bins; and 4. Temporal aggregation can improve estimates when noise contributions are reduced much more than signal variations. Even with a combination of these tools, one needs to be aware of the limitations and physical expectations of the measurement systems.

SECTION 6. Analysis Outline

When faced with a pair of match-up data sets, $\{A_i\}$ and $\{B_i\}$, how should one proceed? Here is an outline for a basic set of calculations, analysis, questions and plots:

1. Compute the averages, $\overline{A}_i$ and $\overline{B}_i$, center of mass, $(\overline{A}_i,\overline{B}_i)$, and relative bias, $[2*(\overline{A}_i - \overline{B}_i)/(\overline{A}_i + \overline{B}_i)]$ for the two sets.

Is the relative bias between the two sets within expectations?

If the data are expected to agree for zero truth, then this can be a first indication of non-unit slope between the two measurements. Are errors in the measurements expected to produce relative errors or offsets?

2. Compute variances and covariances including: $VAR(A_i)$, $VAR(B_i)$, $VAR(A_i–B_i)$ and $COV(A_i,B_i)$.

Is the $VAR(A_i–B_i)$ small relative to the other quantities?

$VAR(A_i–B_i) \approx (\mu_A–\mu_B)^2 VAR(R_i) + \sigma_\varepsilon^2 + \sigma_\delta^2$

This quantity is an approximate upper bound on sum of measurement and match-up noise variabilities.

Are the $VAR(A_i)$ and the $VAR(B_i)$ similar in size?

If not, are there expectations that the larger one is more responsive to real variations?

Is the $COV(A_i,B_i)$ larger than either $VAR(A_i)$ or $VAR(B_i)$?

If it is, then whichever one is smaller has less response to real variability.

If it is smaller than both, then compute the differences.

Do these differences coincide with noise size expectations for the two sets?

3. Compute the Pearson linear correlation coefficient and the slope estimates:

$\rho(A_i,B_i) = COV(A_i,B_i)/SQRT[VAR(B_i)*VAR(A_i)]$

$m_{AB} = COV(A_i,B_i)/VAR(B_i)$

$m_{BA} = COV(A_i,B_i)/VAR(A_i)$

Is $\rho(A_i,B_i)$ close to 1?

The difference between 1 and the absolute value of this quantity tells us how much relative noise is in the measurements. (If ρ is negative, then you are having a problem with anticorrelation – negative slopes – and there is a fundamental disconnect in your model assumptions.)

Do you believe that one measurement is much less noisy than the other?

If so, then slope estimates using it as the independent variable should be closer to the correct answers.

Do you have a good estimate of the noise in one of the measurements?

Consider correcting the slope estimate obtained with that measurement as the independent variable. Recall that $\mu_A/\mu_B \approx m_{AB} / [1 – \sigma_\delta^2 / VAR(B_i)]$ and $\mu_B/\mu_A \approx m_{BA} / [1 – \sigma_\varepsilon^2 / VAR(A_i)]$.

Is the $SQRT(m_{BA}/m_{AB})$ close to 1?

If you believe that the noise is creating similar problems for both measurements, then this quantity is a reasonable value for the slope of B versus A.

Does the interval $[m_{BA},1/m_{AB}]$ contain a reasonable value for the slope?

This may be the best you can do, and even then you should extend this interval with the standard errors for the two slope estimates.

4. Create a scatter plot of $\{B_i\}$ versus $\{A_i\}$.

Overplot three straight lines through the center of mass, one with a slope of m_{BA}, one with a slope of $1/m_{AB}$, and one with either a slope of 1 or a slope of $sqrt(m_{BA}/m_{AB})$.

5. Create a plot of differences in binned values using each of the individual measurements and the sums of the measurements divided by two as three different binning variables.

Does the binning using the sums show any systematic deviations from B-A=0?

6. Sort each data set from smallest to largest and plot the differences as a function of this order.

Does this plot show any systematic deviations from a horizontal line?

7. Consider temporal aggregation to reduce the noise while maintaining the truth variability and repeat the above exercises noting the changes.
Examine time series of the differences between the match-up slope-adjusted data for trends.

8. Collect the above results and consider them together with information on the measurements, retrieval algorithms, match-up process and real world variability.

9. Consider using more advanced analysis techniques for time series regression or retrieval algorithms.

ACKNOWLEDGMENTS

I would like to thank V. Fioletov and D. Crosby for their related work which has helped me to sort out pieces of the statistical puzzle presented here, E. Beach for providing the match-up data sets, T. Beck and Z. Zhang for providing details on the satellite retrievals, L. Crone, R. Mundakkara, and J. Sullivan for reviewing the report, and Y. Duboscq for proofing the manuscript. I would also like to thank the Mauna Loa Dobson Station team for providing a high-quality data record.

BIBLIOGRAPHY

Neter, J., and W. Wasserman, Applied Linear Statistical Models, Richard D. Irwin, Inc., Homewood IL USA, 1974.

Press, W.H., S.A., Teulkolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes in FORTRAN 77: The Art of Scientific Computing, Cambridge University Press, Cambridge, UK, 1992. (See Section 15.3.)

Rodgers, C.D., *Characterization and error analysis of profiles retrieved from remote sounding measurements*, JGR **95** (D5), pg. 5587-95, 1990.

WEB RESOURCESES

Version 8 SBUV/2 ozone retrieval algorithm documents, calibration and validation information, and data are available through links at
http://www.orbit.nesdis.noaa.gov/smcd/spb/ozone/pubs_docs.html
http://www.orbit.nesdis.noaa.gov/smcd/spb/calibration/icvs/sbuvdoc.html
http://www2.ncdc.noaa.gov/docs/klm/
http://www.orbit.nesdis.noaa.gov/pub/smcd/spb/ozone/dvd_v8/DVDhtml/V8_Algorithm_Description.html
http://toms.gsfc.nasa.gov/version8/v8toms_atbd.pdf
and
http://disc.sci.gsfc.nasa.gov/data/datapool/TOMS/DVD-ROMs/

The Dobson Instrument and algorithm are described at
http://www.mlo.noaa.gov/programs/esrl/dobson/dobson.html
and
http://www.srrb.noaa.gov/research/umkehr/

APPENDIX 1. Rotated Coordinates

One approach to the intercomparison problem is to work in a transformed coordinate system. If we choose a coordinate system rotated 45° with respect to the original A/B system, then the line A=B becomes one of the axes. Specifically, define

$$U_i = A_i + B_i = (\mu_A + \mu_B)\, R_i + a + b + \varepsilon_i + \delta_i \quad \text{and} \quad U = A + B = (\mu_A + \mu_B)\, R + a + b$$

and

$$V_i = A_i - B_i = (\mu_A - \mu_B)\, R_i + a - b + \varepsilon_i - \delta_i \quad \text{and } V = A - B = (\mu_A - \mu_B)\, R + a - b$$

$$V = (\mu_A - \mu_B)/(\mu_A + \mu_B)\, U + e$$

$$V = \mu_{VU}\, U + e$$

$$V_i = (\mu_A - \mu_B)/(\mu_A+\mu_B)\, (U_i - \varepsilon_i - \delta_i)+e+\varepsilon_i - \delta_i \qquad V = (\mu_A/\mu_B - 1)/(\mu_A/\mu_B + 1)\, U + e$$

If $\mu_A{-}\mu_B$ is close to zero, then the V_i values should be small. If we fit $\{V_i\}$ with $\{U_i\}$, then we find

$$\begin{aligned} VAR(U_i) &= VAR(A_i) + 2\, COV(A_i,B_i) + VAR(B_i) \\ &= (\mu_A + \mu_B)^2\, VAR(R_i) + (\mu_A + \mu_B)COV(R_i,\varepsilon_i+\delta_i) + VAR(\varepsilon_i+\delta_i) \\ &\approx (\mu_A + \mu_B)^2\, VAR(R_i) + VAR(\varepsilon_i) + VAR(\delta_i) \end{aligned}$$

and

$$\begin{aligned} COV(V_i,U_i) &= VAR(A_i) - VAR(B_i) \\ &= (\mu_A^2-\mu_B^2)VAR(R_i) + 2\mu_A COV(R_i,\varepsilon_i) - 2\mu_B COV(R_i,\delta_i) + VAR(\varepsilon_i) - VAR(\delta_i) \\ &\approx (\mu_A^2-\mu_B^2)VAR(R_i) + VAR(\varepsilon_i) - VAR(\delta_i) \end{aligned}$$

So

$$m_{VU} = COV(V_i,U_i)/VAR(U_i)$$

and substituting for V_i in terms of U_i gives

$$\begin{aligned} m_{VU} &= (\mu_A - \mu_B)/(\mu_A + \mu_B)\, [1 - COV(\varepsilon_i+\delta_i,U_i)/VAR(U_i)] + COV(\varepsilon_i-\delta_i,U_i)/VAR(U_i) \\ &\approx (\mu_A - \mu_B)/(\mu_A + \mu_B)\, \{1 - [VAR(\varepsilon_i)+VAR(\delta_i)]/VAR(U_i)\} + [VAR(\varepsilon_i) - VAR(\delta_i)]/VAR(U_i) \end{aligned}$$

Notice the following: 1) The slope estimate m_{VU} equals 0 if the $VAR(A_i)$ equals the $VAR(B_i)$ [that is, if the $VAR(A_i)/VAR(B_i)$ equals 1]; 2)The bias in the first term involves the sum of the variances of the noise terms; and 3) The second term in the approximation for m_{VU} is small if σ_ε^2 and σ_δ^2 are similar in size. The good news is that the bias in absolute value for the estimate introduced by the noise variance term is to a quantity we expect to be small, namely, $(\mu_A - \mu_B)/(\mu_A + \mu_B)$. The neutral news is that the accuracy of this estimate also relies on a second term which is only small for similar size noise as seen in the earlier analysis. If we ignore the second term and convert this estimate to one for $\mu_A/\mu_B \approx (1+m_{VU})/(1-m_{VU})$, then the estimate is close to the $SQRT[VAR(A_i)/VAR(B_i)]$ result. For small m_{VU}, we have

$$(1+m_{VU})/(1-m_{VU}) \approx 1+2{*}m_{VU}$$

Notice that

$$2{*}[VAR(\varepsilon_i)+VAR(\delta_i)]/VAR(U_i) \approx [VAR(\varepsilon_i)/VAR(A_i) + VAR(\delta_i)/VAR(B_i)]/2$$

for A and B of similar size variances. So our slope error is approximately the same proportion as the error average for the non-rotated case but multiplied by $(\mu_A-\mu_B)/(\mu_A+\mu_B)$. This good result only takes place if the quantity given by $COV(\varepsilon_i-\delta_i,U_i)/VAR(U_i)$ is small.

APPENDIX 2. Histograms.

Not one of my favorites but for completeness, Figure 8 gives the histograms for the four cases with 3-DU wide intervals to get the counts. The intervals began and ended at ½ DU values to avoid problems from the unit precision of the Dobson station values. The data were not adjusted for the biases between the sets. Such adjustments would improve the appearance of the results for the first and fourth cases.

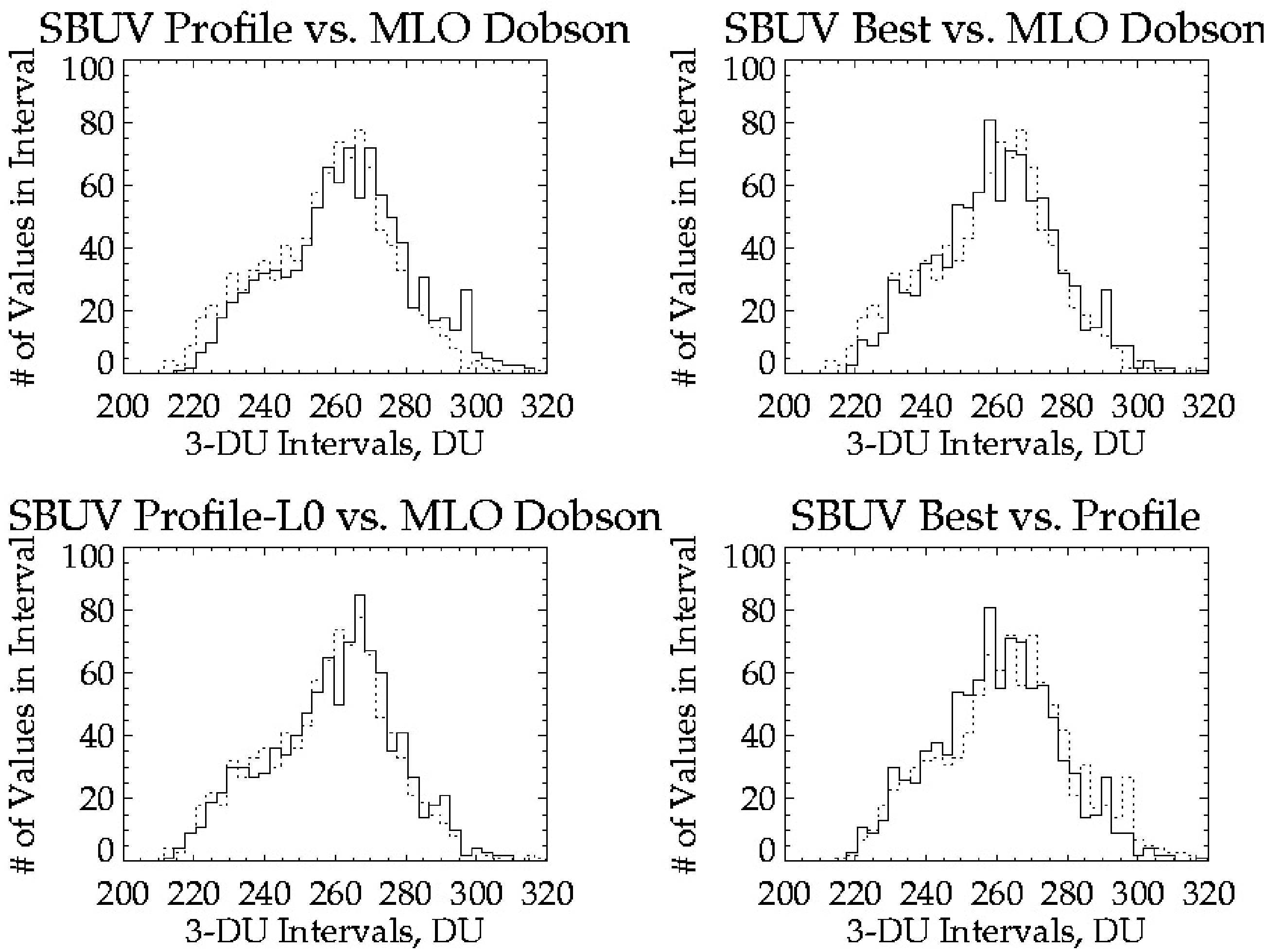


Figure 8. Histograms of data values for the four cases with 3-DU intervals. Cases are arranged as in earlier Figures. The dotted lines are for Dobson (G), Dobson (G), Dobson (G) and Profile (Sp). The solid lines are for Profile (Sp), Best (Sb), Profile Adjusted (Spa) and Best (Sb).

Sometimes representing the data by using histograms such as these can show evidence of differences in the two systems from a new point of view. Fitting with a distribution, *e.g.*, the normal distribution, often just recovers the mean and variance information. These plots can be more helpful if one has lots and lots of data or can stratify the results. If one knows the truth values, then a histogram of the differences can aid in investigating the retrieval or measurement errors. It is interesting to see how the adjustment between the top left and bottom left for the Sp data produces much better agreement in the distributions with G, and how even though the variance of the differences is small for Sb versus Sp, the histogram for the fourth case (lower right) does not look that impressive. One could compare the histograms for two data sets with each set normalized by its mean and variance.

APPENDIX 3. Interactive Data Language (IDL) Code to Generate Synthetic Data Sets

```
        ; INITIALIZE RANDOM NUMBER SEED
ISEED=100675
N=401
        ; CREATE TRUTH SET
R=3+2*DINDGEN(N)/(N-1.)
SAVD=FLTARR(3,5000)
;;;; II=0 REMOVE LOOP FOR SAMPLE CASE
;;;; 500 REPETITIONS OF THE EXPERIMENT
FOR II=0,499 DO BEGIN
        ; CREATE NOISY A AND B SETS
  A=R+RANDOMN(ISEED,N)/3.
  TEMP=RANDOMN(ISEED,5001)
  B=R+RANDOMN(ISEED,N)/4.
        ; COMPUTE MEANS AND STANDARD DEVIATIONS
  AM=TOTAL(A)/N & BM=TOTAL(B)/N & RM=TOTAL(R)/N
  ASD=SQRT(TOTAL((A-AM)^2)/(N-1.))
  BSD=SQRT(TOTAL((B-BM)^2)/(N-1.))
  RSD=SQRT(TOTAL((R-RM)^2)/N)
  COVAB=TOTAL((A-AM)*(B-BM)^2)
  RHO=CORRELATE(A,B)
  AMBSD=SQRT(TOTAL((A-AM-B+BM)^2)/(N-1.))
  APBSD=SQRT(TOTAL((A-AM+B-BM)^2)/(N-1.))
  AMRSD=SQRT(TOTAL((A-R)^2)/N)
  BMRSD=SQRT(TOTAL((B-R)^2)/N)
        ; COMPUTE SUBSECTION 1.1 ESTIMATES
  RESD=SQRT(APBSD2-AMBSD2)/2.
  AMRESD=SQRT(ASD2-RESD2)
  BMRESD=SQRT(BSD2-RESD2)
  SAVD(0,II)=RESD2
  SAVD(1,II)=AMRESD2
  SAVD(2,II)=BMRESD2
ENDFOR
AS=FLTARR(3)
FOR I=0,2 DO AS(I)=TOTAL(SAVD(I,*))/5000.
SS=AS
FOR I=0,2 DO SS(I)=SQRT(TOTAL((SAVD(I,*)-AS(I))^2)/4999.)
PRINT, RM,AM,BM, AM-RM,BM-RM
PRINT, RSD2,RESD2, RSD,RESD
PRINT, ASD2,RSD2+1./9.,ASD,SQRT(RSD2+1./9.)
PRINT, BSD2,RSD2+1./16.,BSD,SQRT(BSD2+1/16.)
PRINT, AMRSD2,AMRESD2, 1./9., AMRSD,AMRESD,1./3.
PRINT, BMRSD2,BMRESD2, 1./16., BMRSD,BMRESD,1./4.
         ; SAMPLE CASE SECTION
         ; SORT DATA
SA=SORT(A) & SB=SORT(B) & SAB=SORT(A+B) & SR=SORT(R)
NS=40
         ; BIN DATA IN 10-PERCENTILES
BA=FLTARR(3,10) & BB=BA & BAB=BA & BR=BA
FOR I=0,9 DO BEGIN
  IA=SA(I*NS:I*NS+NS-1)
  BA(0,I)=TOTAL(A(IA)) & BA(1,I)=TOTAL(B(IA)) & BA(2,I)=TOTAL(R(IA))
  IB=SB(I*NS:I*NS+NS-1)
  BB(0,I)=TOTAL(A(IB)) & BB(1,I)=TOTAL(B(IB)) & BB(2,I)=TOTAL(R(IB))
  IAB=SAB(I*NS:I*NS+NS-1)
  BAB(0,I)=TOTAL(A(IAB)) & BAB(1,I)=TOTAL(B(IAB)) &
```

```
  BAB(2,I)=TOTAL(R(IAB))
  IR=INDGEN(NS)+NS*I
  BR(0,I)=TOTAL(A(IR)) & BR(1,I)=TOTAL(B(IR)) & BR(2,I)=TOTAL(R(IR))
ENDFOR
BA=BA/NS & BB=BB/NS & BAB=BAB/NS & BR=BR/NS
         ; PLOT RESULTS
RAN=[-1,3]+3.
!P.TITLE='NOISY DATA EXPERIEMENTS WITH RANDOM NUMBER DATA
!X.TITLE='DATA WITHOUT NOISE'
!Y.TITLE='DATA WITH N(0,1/3) NOISE'
PLOT, R,A,PSYM=2,XRANGE=RAN,YRANGE=RAN
         ; COMPUTE LINEAR FITS
CCRA=POLY_FIT(R,A,1,RAFIT)
CCAR=POLY_FIT(A,R,1,ARFIT)
CCRB=POLY_FIT(R,B,1,RBFIT)
CCBR=POLY_FIT(B,R,1,BRFIT)
CCAB=POLY_FIT(A,B,1,ABFIT)
CCBA=POLY_FIT(B,A,1,BAFIT)
OPLOT, RAN,RAN
OPLOT, RAN,CCRA(0)+CCRA(1)*RAN,LINESTYLE=1
OPLOT, CCAR(0)+CCAR(1)*RAN,RAN,LINESTYLE=2
PRINT, CCRA,CCAR,CCRA(1),1./CCAR(1),CCRB(1),1/CCBR(1)
READ,ABC ; PAUSE TO LOOK AT THE RESULTS
!P.TITLE='NOISY DATA EXPERIEMENTS WITH RANDOM NUMBER DATA
!X.TITLE='DATA WITH N(0,1/3) NOISE'
!Y.TITLE='DATA WITH N(0,1/4) NOISE'
PLOT,A,B,PSYM=1,XRANGE=RAN,YRANGE=RAN,SYMSIZE=.5
OPLOT, RAN,RAN
OPLOT, RAN,CCAB(0)+CCAB(1)*RAN,LINESTYLE=1
OPLOT, CCBA(0)+CCBA(1)*RAN,RAN,LINESTYLE=2
OPLOT, BA(0,*),BA(1,*),PSYM=2,SYMSIZE=2
OPLOT, BAB(0,*),BAB(1,*),PSYM=4,SYMSIZE=2
OPLOT, BB(0,*),BB(1,*),PSYM=5,SYMSIZE=2
PRINT, CCAB(1),1./CCBA(1)
PRINT,(1.+BMRESD2/RESD2),CCAB(1)*(1.+AMRESD2/RESD2)
PRINT,(1.+AMRSD2/RSD2)*CCAB(1),CCBA(1)*(1.+BMRSD2/RSD2)
STOP
END
```

NESDIS 104 Summary of the NOAA/NESDIS Workshop on Development of a Coordinated Coral Reef Research and Monitoring Program. Jill E. Meyer and H. Lee Dantzler, August 2001.

NESDIS 105 Validation of SSM/I and AMSU Derived Tropical Rainfall Potential (TRaP) During the 2001 Atlantic Hurricane Season. Ralph Ferraro, Paul Pellegrino, Sheldon Kusselson, Michael Turk, and Stan Kidder, August 2002.

NESDIS 106 Calibration of the Advanced Microwave Sounding Unit-A Radiometers for NOAA-N and NOAA-N=. Tsan Mo, September 2002.

NESDIS 107 NOAA Operational Sounding Products for Advanced-TOVS: 2002. Anthony L. Reale, Michael W. Chalfant, Americo S. Allergrino, Franklin H. Tilley, Michael P. Ferguson, and Michael E. Pettey, December 2002.

NESDIS 108 Analytic Formulas for the Aliasing of Sea Level Sampled by a Single Exact-Repeat Altimetric Satellite or a Coordinated Constellation of Satellites. Chang-Kou Tai, November 2002.

NESDIS 109 Description of the System to Nowcast Salinity, Temperature and Sea nettle (*Chrysaora quinquecirrha*) Presence in Chesapeake Bay Using the Curvilinear Hydrodynamics in3-Dimensions (CH3D) Model. Zhen Li, Thomas F. Gross, and Christopher W. Brown, December 2002.

NESDIS 110 An Algorithm for Correction of Navigation Errors in AMSU-A Data. Seiichiro Kigawa and Michael P. Weinreb, December 2002.

NESDIS 111 An Algorithm for Correction of Lunar Contamination in AMSU-A Data. Seiichiro Kigawa and Tsan Mo, December 2002.

NESDIS 112 Sampling Errors of the Global Mean Sea Level Derived from Topex/Poseidon Altimetry. Chang-Kou Tai and Carl Wagner, December 2002.

NESDIS 113 Proceedings of the International GODAR Review Meeting: Abstracts. Sponsors: Intergovernmental Oceanographic Commission, U.S. National Oceanic and Atmospheric Administration, and the European Community, May 2003.

NESDIS 114 Satellite Rainfall Estimation Over South America: Evaluation of Two Major Events. Daniel A. Vila, Roderick A. Scofield, Robert J. Kuligowski, and J. Clay Davenport, May 2003.

NESDIS 115 Imager and Sounder Radiance and Product Validations for the GOES-12 Science Test. Donald W. Hillger, Timothy J. Schmit, and Jamie M. Daniels, September 2003.

NESDIS 116 Microwave Humidity Sounder Calibration Algorithm. Tsan Mo and Kenneth Jarva, October 2004.

NESDIS 117 Building Profile Plankton Databases for Climate and EcoSystem Research. Sydney Levitus, Satoshi Sato, Catherine Maillard, Nick Mikhailov, Pat Cadwell, Harry Dooley, June 2005.

NESDIS 118 Simultaneous Nadir Overpasses for NOAA-6 to NOAA-17 Satellites from 1980 and 2003 for the Intersatellite Calibration of Radiometers. Changyong Cao, Pubu Ciren, August 2005.

NESDIS 119 Calibration and Validation of NOAA 18 Instruments. Fuzhong Wang and Tsan Mo, December 2005.

NESDIS 120 The NOAA/NESDIS/ORA Windsat Calibration/Validation Collocation Database. Laurence Connor, February 2006.

NESDIS 121 Calibration of the Advanced Microwave Sounding Unit-A Radiometer for METOP-A. Tsan Mo, August 2006.

NESDIS 122 JCSDA Community Radiative Transfer Model (CRTM). Yong Han, Paul van Delst, Quanhua Liu, Fuzhong Weng, Banghua Yan, Russ Treadon, and John Derber, December 2005.

NESDIS 123 Comparing Two Sets of Noisy Measurements. Lawrence E. Flynn, April 2007

NOAA SCIENTIFIC AND TECHNICAL PUBLICATIONS

The National Oceanic and Atmospheric Administration **was established as part of the Department of Commerce on October 3, 1970. The mission responsibilities of NOAA are to assess the socioeconomic impact of natural and technological changes in the environment and to monitor and predict the state of the solid Earth, the oceans and their living resources, the atmosphere, and the space environment of the Earth.**

The major components of NOAA regularly produce various types of scientific and technical information in the following types of publications

PROFESSIONAL PAPERS – Important definitive research results, major techniques, and special investigations.

CONTRACT AND GRANT REPORTS – Reports prepared by contractors or grantees under NOAA sponsorship.

ATLAS – Presentation of analyzed data generally in the form of maps showing distribution of rainfall, chemical and physical conditions of oceans and atmosphere, distribution of fishes and marine mammals, ionospheric conditions, etc.

TECHNICAL SERVICE PUBLICATIONS – Reports containing data, observations, instructions, etc. A partial listing includes data serials; prediction and outlook periodicals; technical manuals, training papers, planning reports, and information serials; and miscellaneous technical publications.

TECHNICAL REPORTS – Journal quality with extensive details, mathematical developments, or data listings.

TECHNICAL MEMORANDUMS – Reports of preliminary, partial, or negative research or technology results, interim instructions, and the like.

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service
Washington, D.C. 20233