JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. D14, PAGES 17,149–17,161, JULY 27, 1998

Factors affecting the detection of trends: Statistical considerations and applications to environmental data

Elizabeth C. Weatherhead,[1] Gregory C. Reinsel,[2] George C. Tiao,[3] Xiao-Li Meng,[4] Dongseok Choi,[4] Wai-Kwong Cheang,[2] Teddie Keller,[5] John DeLuisi,[6] Donald J. Wuebbles,[7] James B. Kerr,[8] Alvin J. Miller,[9] Samuel J. Oltmans,[10] and John E. Frederick,[11]

Abstract. Detection of long-term, linear trends is affected by a number of factors, including the size of trend to be detected, the time span of available data, and the magnitude of variability and autocorrelation of the noise in the data. The number of years of data necessary to detect a trend is strongly dependent on, and increases with, the magnitude of variance (σ_N^2) and autocorrelation coefficient (ϕ) of the noise. For a typical range of values of σ_N^2 and ϕ the number of years of data needed to detect a trend of 5%/decade can vary from ~10 to >20 years, implying that in choosing sites to detect trends some locations are likely to be more efficient and cost-effective than others. Additionally, some environmental variables allow for an earlier detection of trends than other variables because of their low variability and autocorrelation. The detection of trends can be confounded when sudden changes occur in the data, such as when an instrument is changed or a volcano erupts. Sudden level shifts in data sets, whether due to artificial sources, such as changes in instrumentation or site location, or natural sources, such as volcanic eruptions or local changes to the environment, can strongly impact the number of years necessary to detect a given trend, increasing the number of years by as much as 50% or more. This paper provides formulae for estimating the number of years necessary to detect trends, along with the estimates of the impact of interventions on trend detection. The uncertainty associated with these estimates is also explored. The results presented are relevant for a variety of practical decisions in managing a monitoring station, such as whether to move an instrument, change monitoring protocols in the middle of a long-term monitoring program, or try to reduce uncertainty in the measurements by improved calibration techniques. The results are also useful for establishing reasonable expectations for trend detection and can be helpful in selecting sites and environmental variables for the detection of trends. An important implication of these results is that it will take several decades of high-quality data to detect the trends likely to occur in nature.

1. Introduction

The impact of human intervention in a changing environment has brought about increased concern for detecting trends in various types of environmental data. A variety of studies have attempted to detect long-term trends in geophysical variables such as atmospheric ozone [*Reinsel et al.*, 1994; *Stolarski et al.*, 1991, 1992], stratospheric temperature [*Miller et al.*, 1992], and ultraviolet (UV) radiation at the Earth's surface [*Scotto et al.*, 1988; *Weatherhead et al.*, 1997]. These studies have revealed that detection of trends is difficult and requires statistical techniques that take into account some of the realistic problems which frequently occur with the measurement of geophysical data. As political decisions and future scientific efforts may be based on the results of environmental trend studies, achieving the most accurate trend estimates in the shortest time period is important to ensuring that appropriate actions are taken. Factors affecting the detection of linear trends are outlined in this paper, and examples of trend detection analyses of data are highlighted. The results show that for most expected environmental changes, several decades of high-quality data will be needed before such changes will be detectable. The results also show that certain decisions regarding site location, instrument maintenance, and calibration can influence the accuracy of trend estimation and hence detection of trends.

Statistical criteria for detecting linear trends were presented by *Tiao et al.* [1990], where it was shown that the precision of

[1]Cooperative Institute for Research in the Environmental Sciences, University of Colorado, Boulder.
[2]Department of Statistics, University of Wisconsin, Madison.
[3]Graduate School of Business, University of Chicago, Chicago, Illinois.
[4]Department of Statistics, University of Chicago, Chicago, Illinois.
[5]National Center for Atmospheric Research, Boulder, Colorado.
[6]NOAA Air Resources Laboratory, Boulder, Colorado.
[7]Department of Atmospheric Sciences, University of Illinois, Champaign.
[8]Atmospheric Environment Service, Downsview, Ontario, Canada.
[9]NOAA National Weather Service, Washington, D. C.
[10]NOAA Climate Monitoring and Diagnostics Laboratory, Boulder, Colorado.
[11]Department of Geophysical Sciences, University of Chicago, Chicago, Illinois.

Paper number 98JD00995.
0148-0227/98/98JD-00995$09.00

trend estimates is strongly influenced by the variability and autocorrelation of the underlying noise process. The precision of a trend estimate in turn directly determines the number of years of data required to detect a trend of given magnitude or the magnitude of trend that can be detected with a given number of years of data. This paper explores more fully, in both a theoretical and an applied sense, how various factors affect the ability to detect a trend through their influence on the precision of the trend estimate. The uncertainty of estimates for trend detectability is also explored. Section 2 of this paper addresses the ability to detect a trend in a single data set and how the number of years of data needed to detect a given trend is dependent on the magnitude of and autocorrelation in the natural variation of the data, as well as how these parameters can vary significantly among environmental parameters as well as from location to location. Section 3 addresses the impact of sudden interventions on the ability to detect trends. Such level shifts have been observed in empirical trend studies for a variety of data sets [e.g., *Reinsel et al.*, 1994; *Weatherhead et al.*, 1997]. The basic statistical theory used in this paper is covered in more detail in standard textbooks on regression and time series analysis [e.g., *Box et al.*, 1994]; however, here applications and examples are presented with special emphasis on the case of linear trend estimation with examples from environmental data sets.

The detectability of a trend can be summarized in several ways. Two common ways are through the precision of a trend estimate, as measured by its standard deviation, and through the number of years of data required to detect a trend of given magnitude using the trend estimate. Which summary is employed is determined by what data are available or are being planned for. If the data have already been collected, generally the former is more useful, while if an experiment is being planned or a monitoring project started, the latter is often more useful. In each of the following sections, the statistical modeling approach will be introduced, and the precision of a trend estimate will be examined. Then the number of years required to detect a given trend will be discussed, followed by examples with actual data sets to illustrate the precision of trend estimates. While examples will be supplied using atmospheric data, the results are general and are applicable to data from a variety of sources. Details of the statistical approaches will be more fully outlined in the Appendix.

2. Basic Trend Evaluation: Effects of autocorrelation and variability on trend estimation and detection

Changes in environmental variables are often statistically modeled as being a smooth linear change. While this may never be the case, it allows a simple approximation of the direction and magnitude of the changes in the data, which may be adequate for many practical purposes. More importantly, results from the linear trend models are commonly used and are familiar to scientists and policy makers; thus it is important to examine various issues under such models. The linear trend in environmental data is often expressed in percent change per decade. There is often a great deal of variation on different timescales occurring within the data in addition to the linear trend. For independent (uncorrelated) time series data the length of time to detect the trend is determined by the magnitude of variation of the noise. However, environmental data are often autocorrelated. For instance, higher than normal temperature on one day is often associated with higher than normal temperature on the next day. This positive autocorrelation tends to confound with a linear trend and therefore increases the length of time required to detect a given trend.

2.1. Basic Statistical Modeling

In many statistical trend analyses of monthly time series data on a geophysical variable, such as total ozone, stratospheric temperature, or UV radiation, a model of the following form has frequently been found to be useful. Let Y_t be our measurement on the geophysical variable of interest, either on its original scale or after a transformation (most often the log transformation, e.g., log UV). Then the linear trend model assumes $Y_t = \mu + S_t + \omega X_t + N_t$, where μ is a constant term, $X_t = t/12$ represents the linear trend function, and ω is the magnitude of the trend per year. S_t is a seasonal component which can often be represented as $S_t = \Sigma_{j=1}^{4} [\beta_{1,j} \sin(2\pi jt/12) + \beta_{2,j} \cos(2\pi jt/12)]$. While the seasonal component is essential in practical modeling of geophysical time series, estimation of this component does not have much impact on the statistical properties of the estimates of the other terms in the model. Hence the seasonal component will not be included in our subsequent statistical derivations, for convenience, although it will be included in the analyses for the empirical data applications. Therefore to investigate the effects of magnitude and autocorrelation of noise on trend estimates, we consider a simple trend model of the form

$$Y_t = \mu + \omega X_t + N_t, \qquad t = 1, \cdots, T. \tag{1}$$

For the unexplained portion of the data the noise N_t is assumed to be autoregressive of the order of 1 [AR(1)]; that is, $N_t = \phi N_{t-1} + \varepsilon_t$, where the ε_t are independent random variables with mean zero and common variance σ_ε^2. It will also be assumed that $-1 < \phi < 1$, so the noise process $\{N_t\}$ is stationary. This model allows for the noise to be (auto)correlated among successive measurements, such that $\phi = \text{Corr}(N_t, N_{t-1})$. The autocorrelation, which is typically positive, may be the result of various natural factors which give rise to somewhat smoothly varying changes in N_t over time. Such natural factors may not always be known or measurable, and the lagged value N_{t-1} in the model can sometimes be regarded as a proxy to represent these natural factors which dynamically influence the current noise value N_t. It is also noted that the variance of the noise N_t is directly related to the variance σ_ε^2 of the white noise process $\{\varepsilon_t\}$ in this model by $\sigma_N^2 = \text{Var}(N_t) = \sigma_\varepsilon^2/(1 - \phi^2)$.

2.2. Trend Estimation

A common situation of trend evaluation is when a given number of years of data have been collected and one wants to estimate a trend based on the existing data. For such a situation, one wants to estimate the trend as well as the precision (standard deviation) of the trend estimate. Let $\hat{\omega}$ denote the generalized least squares (GLS) estimator of the trend ω in model (1), and let $\sigma_{\hat{\omega}}$ = s.d. $(\hat{\omega})$ denote the corresponding standard deviation of $\hat{\omega}$. The exact form of $\sigma_{\hat{\omega}}$ is derived in the Appendix, with $\sigma_{\hat{\omega}}^2 = \text{Var}(\hat{\omega})$ given in (A5). In the Appendix it is also shown that a useful and quite accurate approximation is

$$\sigma_{\hat{\omega}} \approx \frac{\sigma_\varepsilon}{(1-\phi)} \frac{1}{n^{3/2}} = \frac{\sigma_N}{n^{3/2}} \sqrt{\frac{1+\phi}{1-\phi}} \tag{2}$$

where $n = T/12$ denotes the number of years of data.

As seen from approximation (2), the precision of $\hat{\omega}$ is directly a function of the magnitude of variation and the autocorrelation of the noise, as well as the fixed number of years of available data. Specifically, data with larger variance and higher positive autocorrelation of the noise decrease the precision (i.e., increase s.d. ($\hat{\omega}$)) of the trend estimate. Large σ_N means large noise in the data, making any signal or trend more difficult to detect, the standard "signal-to-noise" issue. Increases in positive autocorrelation (ϕ) tend to increase the length of trend-like segments in the data and thus make the estimation of the real trend more difficult.

Consequently, failure in taking into account the proper variance and autocorrelation of the noise can lead to a false level of precision in the trend estimate, in that the assumed standard deviation of the trend estimate will substantially understate the actual uncertainty. For instance, a trend result may be obtained for typical autocorrelated environmental time series data using a statistical model that ignores autocorrelation, for example, assuming $\phi = 0$ in model (1). The result obtained might falsely indicate a statistically significant trend at the 95% confidence level, whereas if the appropriate statistical model were used which does not ignore the autocorrelation, the actual precision of the trend estimate might be found to be substantially less.

For any given value of ϕ, we see from the approximation in (2) that $\sigma_{\hat{\omega}}/\sigma_\varepsilon$ decreases as n increases at a rate essentially proportional to $n^{-3/2}$. The standard deviation of $\hat{\omega}$ increases as ϕ increases, and with a fixed value of σ_ε, the standard deviation of $\hat{\omega}$ for the moderate autocorrelation of $\phi = 0.5$ is approximately doubled relative to its value when $\phi = 0$. However, it must be noted that AR(1) noise processes $\{N_t\}$ with the same σ_ε = s.d. (ε_t) and different ϕ values have different variances, since $\sigma_N^2 = \mathrm{Var}(N_t) = \sigma_\varepsilon^2/(1 - \phi^2)$. Hence in terms of σ_N the standard deviation of $\hat{\omega}$ can be written as $\sigma_{\hat{\omega}} = \sigma_N\sqrt{(1 + \phi)/(1 - \phi)}n^{-3/2}$. Thus if σ_N, rather than σ_ε, is fixed, then the standard deviation of $\hat{\omega}$ when $\phi = 0.5$ is in fact approximately $\sqrt{(1 + 0.5)/(1 - 0.5)} = \sqrt{3} \approx 1.73$ times that when $\phi = 0$.

2.3. Trend Detection

The issue of key interest in trend detection is to determine the number of years of data required to detect a trend of a certain magnitude. We shall adopt the commonly used decision rule that a real trend is indicated, at the 5% significance level or 95% confidence level, when $|\hat{\omega}/\sigma_{\hat{\omega}}| > 2$. It is then established by *Tiao et al.* [1990] that the number of years n^* of data required to detect a real trend of specified magnitude $|\omega| = |\omega_0|$, with probability 0.90, is

$$n^* \approx \left[\frac{3.3\sigma_\varepsilon}{|\omega_0|(1 - \phi)}\right]^{2/3} = \left[\frac{3.3\sigma_N}{|\omega_0|}\sqrt{\frac{1 + \phi}{1 - \phi}}\right]^{2/3}, \quad (3)$$

where we assume $|\phi| < 1$. We note here that this formula requires σ_N and ω_0 be on the same scale; that is, if ω_0 is in percentage, then σ_N must be converted to percentage as well (e.g., dividing the original σ_N by a mean). Also, the value 3.3 in (3) would be larger (smaller) if we require a higher (lower) degree of certainty than 90%.

From this result we can see the importance of both the magnitude of variation and the autocorrelation of the noise on the detection of trend. In general, from (3) we find that the presence of positive autocorrelation ($\phi > 0$) increases the required number of years for trend detection by the factor of $[(1 + \phi)/(1 - \phi)]^{1/3}$, approximately, compared to the case of no autocorrelation (assuming a common value of the noise variance σ_N^2). For numerical illustration, suppose that one is interested in being able to detect a trend of magnitude 5% per decade (ω_0 = 0.5% per year). Table 1 (top) shows the number of years of data that are needed to detect such a trend given data with various values of both the magnitude of variation σ_N and autocorrelation ϕ. The results presented show a variety of features. First, the number of years to detect a given trend is strongly influenced by both the autocorrelation and the variance of noise in the data. For example, a month-to-month variability of 10% ($\sigma_N = 10$) and moderate autocorrelation of $\phi = 0.5$ results in a situation requiring 23.4 years of data to detect a 5% per decade trend. For the same value of σ_N but with zero autocorrelation, $\phi = 0$, the required number of years will be 16.3, a difference of about 7 years compared to the case $\phi = 0.5$. This result also shows the effect of ignoring the presence of autocorrelation in the data (e.g., assume $\phi = 0$ when in fact $\phi = 0.5$). There would then be a false sense of confidence in the ability to detect a trend; that is, the number of years to detect a trend would be underestimated by 7 years.

Note that the results in Table 1 (top) are also presented in terms of the ratio σ_N/ω_0 and can be used for general values of trend (ω_0). For illustration, if it is desired to detect a trend of magnitude ω_0 = 0.25% per year, and the underlying noise process has σ_N = 3.0%, then we can obtain the required number of years from the row in Table 1 (top) with σ_N/ω_0 = 12 (e.g., $n^* = 16.6$ when $\phi = 0.5$).

For further numerical comparison, Table 1 (bottom) shows the number of years n^* of data required for detection of a trend of magnitude 1% per decade; that is, ω_0 = 0.1% per year. Even for small autocorrelation and low variability, it would often require a prohibitively long period of time to detect such a small trend. The long times required for trend detection arise because of the random variability in the data sets. This variability is itself of great scientific interest. For many environmental parameters, for example, height of sea level, precipitation, or UV radiation, a year of unusually high levels once in a while might be of greater relevance than a small upward trend over many years.

2.4. Uncertainties Due To Unknown Variance and Autocorrelation

The approximation for n^* given in (3) not only assumes the linear model (1) is adequate but also that we know the true values of σ_N and ϕ. In practice, even if we assume the adequacy of model (1), we still do not know σ_N or ϕ, and thus we have to estimate them from available data (and prior information); that is, we will use the estimated values ($\hat{\sigma}_\varepsilon$ and $\hat{\phi}$) in place of σ_ε and ϕ, respectively, when we apply (3), resulting in an estimated n^*, $\hat{n}^*$. Given that n^* strongly depends on σ_ε and ϕ, it is natural to worry about the uncertainty in $\hat{n}^*$, particularly if there is large uncertainty in $\hat{\sigma}_\varepsilon$ or in $\hat{\phi}$. In practice, only a few years of data are necessary to estimate σ_ε fairly well, and thus it is often acceptable to ignore the uncertainty in $\hat{\sigma}_\varepsilon$ (i.e., we can treat $\sigma_\varepsilon = \hat{\sigma}_\varepsilon$); significantly longer time is needed to adequately estimate ϕ. Consequently, it is important to assess the uncertainty in $\hat{n}^*$ due to the uncertainty in $\hat{\phi}$, considering that the amount of data available for estimating ϕ at the planning stage (i.e., when we need to estimate the number of years) is typically not large.

The following method provides an approximate 95% confidence interval for n^* when $\hat{\phi}$ is the least squares estimate of ϕ based on M months of data. The derivation is given in the

Table 1. Number of Years of Monthly Data Needed to Detect, With Probability 0.90, a Trend of 5% Per Decade and 1% Per Decade at a 95% Confidence Level for Selected Values of Autocorrelation (ϕ) and Standard Deviation (σ_N) of the Noise

		Value of ϕ								
σ_N	(σ_N/ω_0)	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	Number of Years to Detect a Trend of 5% Per Decade (ω_0 = *0.5% per year*)									
0.25	(0.5)	1.4	1.5	1.6	1.7	1.8	2.0	2.1	2.3	
0.5	(1)	2.2	2.4	2.5	2.7	2.8	3.0	3.3	3.6	4.0
1	(2)	3.5	3.7	4.0	4.3	4.6	4.9	5.3	5.9	6.7
2	(4)	5.6	6.0	6.4	6.8	7.3	7.9	8.6	9.6	11.0
4	(8)	8.9	9.5	10.1	10.8	11.6	12.6	13.8	15.4	17.8
6	(12)	11.6	12.4	13.3	14.2	15.3	16.6	18.2	20.3	23.5
8	(16)	14.1	15.0	16.1	17.2	18.6	20.1	22.1	24.7	28.6
10	(20)	16.3	17.4	18.7	20.0	21.6	23.4	25.7	28.7	33.3
12	(24)	18.4	19.7	21.1	22.6	24.4	26.4	29.0	32.5	37.7
15	(30)	21.4	22.9	24.5	26.2	28.3	30.7	33.7	37.8	43.8
20	(40)	25.9	27.7	29.6	31.8	34.3	37.2	40.9	45.8	53.3
	Number of Years to Detect a Trend of 1% Per Decade (ω_0 = *0.1% per year*)									
0.25	(2.5)	4.1	4.3	4.6	4.9	5.3	5.7	6.2	6.9	7.8
0.5	(5)	6.5	6.9	7.4	7.9	8.5	9.2	10.0	11.2	12.8
1	(10)	10.3	11.0	11.7	12.6	13.5	14.7	16.1	18.0	20.7
2	(20)	16.3	17.4	18.7	20.0	21.6	23.4	25.7	28.7	33.3
4	(40)	25.9	27.7	29.6	31.8	34.3	37.2	40.9	45.8	53.3
6	(60)	34.0	36.3	38.8	41.7	44.9	48.8	53.7	60.2	70.0
8	(80)	41.2	44.0	47.1	50.5	54.5	59.2	65.1	73.0	84.9
10	(100)	47.8	51.0	54.6	58.6	63.2	68.7	75.6	84.7	98.7
12	(120)	53.9	57.6	61.7	66.2	71.4	77.6	85.4	95.8	111.5
15	(150)	62.6	66.9	71.6	76.8	82.9	90.1	99.1	111.2	129.5
20	(200)	75.8	81.0	86.7	93.1	100.4	109.2	120.1	134.8	157.0

The standard deviation σ_N of the noise is expressed in percent variability in the month-to-month data. Table values can also be used for trend detection for general value of ω_0, in terms of the ratio σ_N/ω_0.

Appendix (section A2). First we calculate an estimated uncertainty factor

$$B = \frac{4}{3\sqrt{M}}\sqrt{\frac{1+\hat{\phi}}{1-\hat{\phi}}}. \tag{4}$$

Then the 95% confidence interval for the number of years is given by $(\hat{n}^* e^{-B}, \hat{n}^* e^{B})$, where $\hat{n}^*$ is calculated according to (3) using $\hat{\phi}$ and $\hat{\sigma}_\varepsilon$ for ϕ and σ_ε, respectively. Note that when $B \leq 0.2$, the interval is essentially given by $\hat{n}^*(1 \pm B)$, thus B is the percentage of uncertainty relative to the point estimate $\hat{n}^*$. For larger B the interval $(\hat{n}^* e^{-B}, \hat{n}^* e^{B})$ is not symmetric about $\hat{n}^*$ but rather skewed toward large values.

To illustrate the calculation of this procedure, suppose $\hat{\sigma}_\varepsilon$ = 3.1%, ω_0 = 0.3% per year and $\hat{\phi} = 0.32$. These values are actually from a total ozone data set collected at Tateno, as reported by *Tiao et al.* [1990], who gave $\hat{n}^* \approx 14$ years; the value from (3) is 13.6. Now suppose the value $\hat{\phi} = 0.32$ was computed on the basis of 2 years of data; that is, $M = 24$. Then the B given by (4) is 0.38, and thus an approximate 95% interval for the number of years is $(13.6e^{-0.38}, 13.6e^{0.38}) \approx$ (9 years, 20 years), which is a rather wide interval. If we have 5 years of data to estimate ϕ (i.e., $M = 60$), then $B = 0.24$, and the interval becomes (10.8, 17.3), or approximately 14 ± 3 years. Having 5 years of data to estimate ϕ during the planning stage is likely to be unusual in practice. Assuming M varies from 24 to 60 months and ϕ varies from 0 to 0.5, the relative uncertainty from the interval estimate varies approximately from 20 to 60%, and thus it is clear that the uncertainty in estimating ϕ cannot be ignored when estimating the number of years.

2.5. Applications

The precision of a long-term linear trend estimate has been shown in sections 2.1 and 2.2 to be highly dependent on the variance (σ_ε^2 or σ_N^2) and autocorrelation (ϕ) of the noise. Realistic estimates of these parameters can allow one to estimate the number of years necessary to detect a trend of given magnitude, or the magnitude of trend that can be detected by a fixed number of years of data. Knowledge of these parameter values (σ_N and ϕ) over a variety of different locations can help one determine the most likely locations from which a trend could first be detected. Such an analysis will allow reasonable estimates of the time necessary to detect a given trend.

To illustrate, UV data from 14 Robertson-Berger (RB) meter stations within the United States were examined to assess the autocorrelation ϕ and the magnitude of variability σ_N for trend detection of UV radiation after seasonal variability has been accounted for [see *Weatherhead et al.*, 1997]. Table 2 lists the 14 stations and gives the estimates of ϕ and σ_N and the corresponding number of years required to detect a trend of 5% per decade. The results are also graphically illustrated in Figure 1, which shows the estimates of ϕ and σ_N along with contour curves indicating the number of years of data needed to detect a 5% per decade trend. The plot shows that the range of autocorrelation and variability observed at these stations translates into a rather wide range of the number of years to detect a given trend. It should be noted that the values of σ_N and ϕ have a latitudinal dependence with lower σ_N and higher ϕ observed nearer the equator (see Table 2).

While such an analysis can be helpful in identifying locations likely to detect a given trend earliest, it should be noted that

Table 2. Estimated Standard Deviation ($\hat{\sigma}_N$) and Autocorrelation ($\hat{\phi}$) of Noise in UV Data From 14 Locations Along With the Estimated Number of Years ($\hat{n}^*$) to Detect Trends

Site	Latitude	$\hat{\phi}$	$\hat{\sigma}_N$ (in %)	Reference Trend, % per decade	$\hat{n}^*$ for Reference Trend	$\hat{n}^*$ for 5% Trend, per decade
Seattle, WA	47.5°N	0.12	13.0	3.1	29	21
Bismarck, ND	46.8°N	0.24	14.9	2.6	37	24
Minneapolis, MN	44.9°N	0.13	11.0	3.0	27	19
Detroit, MI	42.4°N	0.12	11.6	3.2	26	20
Des Moines, IA	41.6°N	0.29	13.5	2.9	35	24
Salt Lake City, UT	40.8°N	−0.15	10.5	2.3	25	15
Philadelphia, PA	39.9°N	0.05	12.3	2.9	28	19
Oakland, CA	37.7°N	0.21	9.3	2.3	30	18
Albuquerque, NM	35.1°N	0.31	7.7	2.1	30	17
Fort Worth, TX	32.8°N	0.23	12.4	1.7	45	22
Tucson, AZ	32.3°N	0.27	7.0	1.9	30	16
El Paso, TX	31.8°N	0.18	7.8	1.7	33	16
Tallahassee, FL	30.4°N	0.04	9.8	1.8	33	17
Mauna Loa, HI	19.5°N	0.36	12.0	0.6	97	24

The uncertainties for the estimated number of years presented in this table are all of the order of ±10–15% because the number of years of data here range between 10 and 18 years.

for any given variable, in this case UV radiation, the magnitude of the expected trend may vary by location as well. Thus a more pointed analysis comparing size of expected trends with the ability of a site to detect that trend is of further value. Often a realistic expected trend may not exist, but a reference trend may be established using the best available information. In practice, the estimated number of years to detect a trend is highly sensitive to the magnitude of the reference trend. In practice, a sensitivity study can show the exact dependence. Columns 6 and 7 of Table 2 illustrate the importance of establishing expectations for trend detection in evaluating various sites for their ability to detect trends. On the basis of past ozone trends, UV radiation is not expected to change as much in the lower-latitude sites as in the higher-latitude sites. In particular, using past ozone data from the total ozone mapping spectrometer on the Nimbus 7 satellite as a reference and assuming that UV radiation will increase approximately 1.9% for each 1% decrease in ozone, we can establish reference trends for UV radiation as given in Table 2 (column 5). Calculating the numbers of years to detect these reference trends, shown in column 6 of Table 2, reveals that data from Detroit and Minneapolis are likely to identify their reference trends several years earlier than Tallahassee, El Paso, or Tucson. This is in contrast to the results presented in column 7 of Table 2 which show that data from sites like Tallahassee, El Paso, and Tucson would be likely to detect a common trend of 5% per decade several years sooner than from sites like Detroit or Minneapolis. A more complete comparison would include the estimate of errors in these numbers.

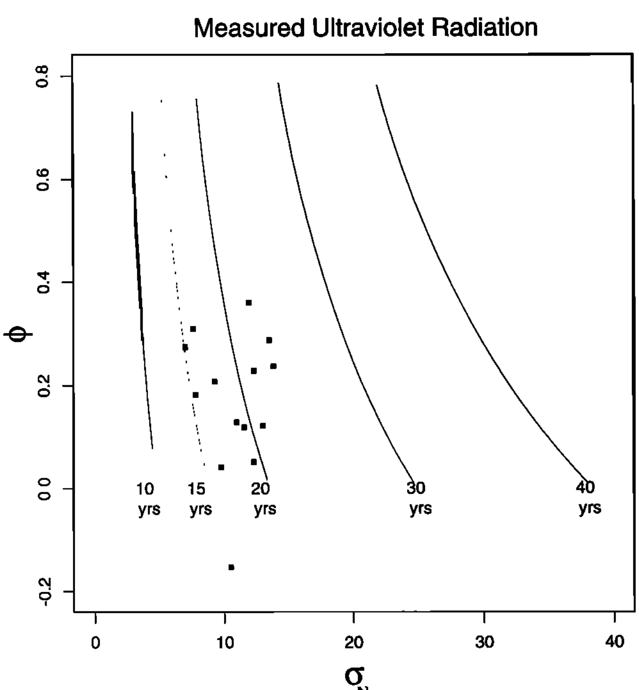

Figure 1. Magnitude (σ_N) and autocorrelation (ϕ) from 14 UV monitoring sites within the United States. Lines show combinations of σ_N and ϕ required to detect a trend of 5% per decade in the stated number of years. This plot shows that some locations appear to be better suited than other locations for detecting a given trend.

In addition to the comparison of detectability at different locations, there is another comparison that is of great interest which this type of analysis can offer. Anthropogenic influences are predicted to create changes in a variety of variables. The variable with the largest expected change may be identified as a key variable to monitor with respect to global change. However, different environmental variables have characteristically different noise associated with them, implying that for particular variables, despite their low expected signal, it may be possible to detect a long-term change relatively early. Figure 2 shows estimates of σ_N in percent of mean value and ϕ for data on six atmospheric variables after seasonal variability has been accounted for, with the data for the first four plots obtained from the same 14 locations in the United States, as given in Table 2. The surface ozone data are from 13 locations internationally, as listed in Table 3. The total ozone data are from satellite measurements over the same locations, as presented in Figure 1 and Table 2. Information on the sources of the data may be found in the Appendix (section A1). The lines in Figure 2 indicate the number of years of data needed to detect a trend of 5% per decade. Figure 2 shows that some environmental parameters appear to be inherently more variable or more

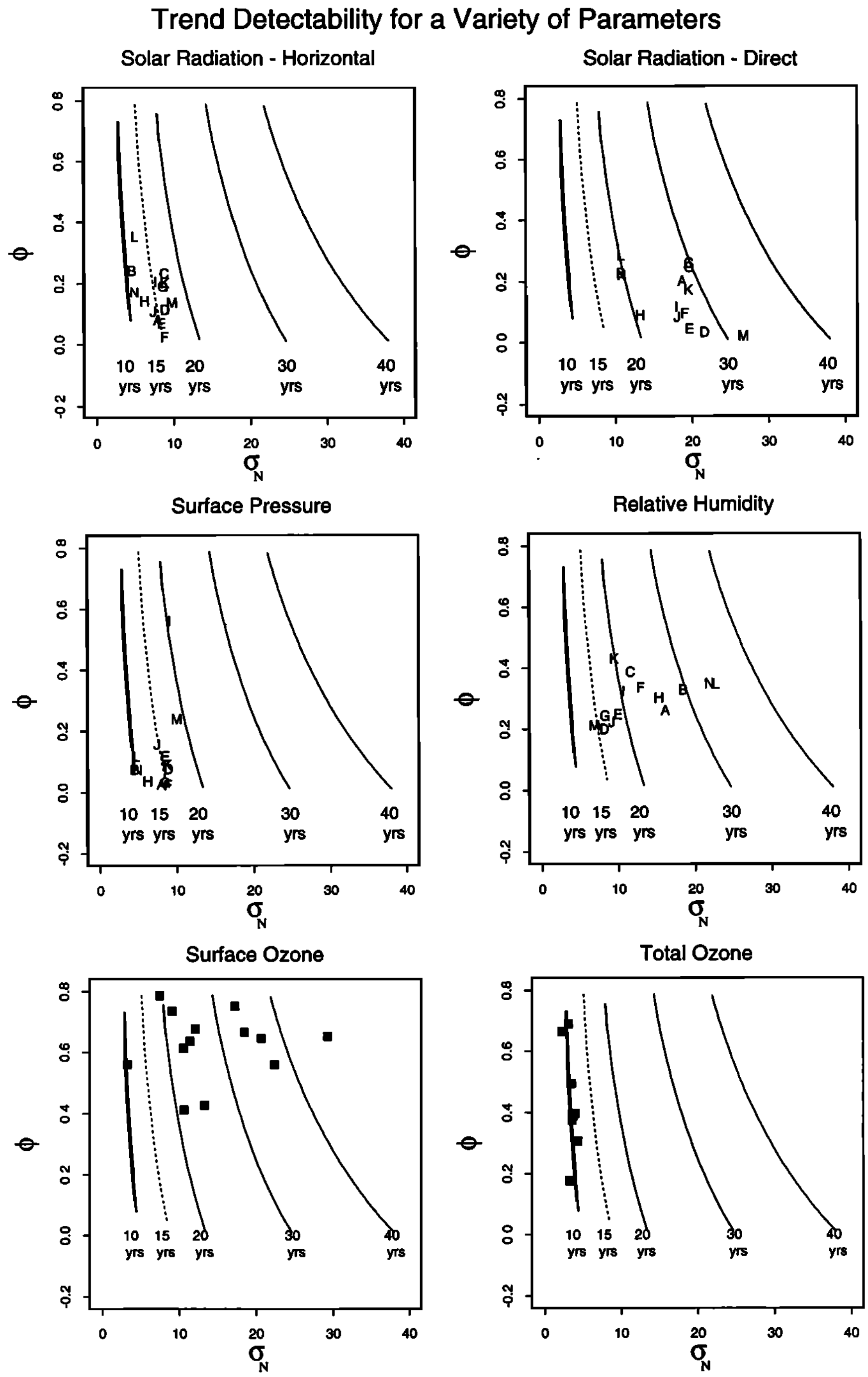

Figure 2. Magnitude (σ_N) and autocorrelation (ϕ) from a variety of different monitoring stations. The first four plots show data from the same 14 locations with the following key for locations: A, Caribou, Maine; B, Bismarck, North Dakota; C, Great Falls, Montana; D, Eugene, Oregon; E, Lander, Wyoming; F, Madison, Wisconsin; G, Omaha, Nebraska; H, Sterling, Virginia; I, Albuquerque, New Mexico; J, Phoenix, Arizona; K, El Paso, Texas; L, Montgomery, Alabama; M, Tallahassee, Florida; N, Brownsville, Texas. (Letter order represents decreasing latitude.) The plot of surface ozone represents data from 13 locations internationally. The plot of total ozone represents data from satellite which correspond to the 14 locations displayed in Figure 1. In all six plots, lines show combinations of σ_N and ϕ required to detect a trend of 5% per decade in the stated number of years. These plots show that some variables appear to be inherently more difficult than others for detecting a fixed trend.

autocorrelated than others. The variables examined for this study indicate that some are capable of showing significant changes sooner than others. Examination of the first four plots which show a number of parameters from the same 14 locations shows that while a site, such as Tallahassee, Florida, may be a preferred site for detecting a trend in one parameter (relative humidity), it may be a less desirable site for detecting a trend in another parameter (solar radiation). This analysis

Table 3. Estimated Standard Deviation ($\hat{\sigma}_N$) and Autocorrelation ($\hat{\phi}$) of Noise in Surface Ozone Data from 13 Locations Along With the Estimated Number of Years ($\hat{n}^*$) to Detect Trends and an Approximate 95% Interval for the Number of Years (in Parenthesis)

Site	M	$\hat{\phi}$	$\hat{\sigma}_N$, %	$\hat{\sigma}_N$, ppb	$\hat{n}^*$ to Detect 5% per Decade Trend and Its Uncertainty	$\hat{n}^*$ to Detect 1 ppb per Decade and Its Uncertainty
Barrow, Alaska	270	0.56	22.5	5.77	−42 (36, 50)	50 (43, 59)
Iceland	39	0.64	11.5	4.35	29 (18, 47)	45 (28, 71)
Ireland	78	0.41	10.7	3.87	23 (17, 29)	34 (26, 43)
Zugspitze	204	0.61	10.6	4.17	27 (22, 33)	43 (35, 52)
Niwot	53	0.74	9.2	4.19	28 (17, 46)	49 (30, 80)
Bermuda	83	0.65	20.7	7.06	44 (32, 61)	63 (45, 87)
Canary	104	0.68	12.1	5.46	32 (23, 43)	55 (40, 75)
Mauna Loa	259	0.67	18.5	6.88	42 (34, 51)	63 (52, 77)
Barbados	53	0.43	13.3	3.68	27 (19, 36)	27 (20, 36)
Samoa	214	0.65	29.4	4.14	56 (45, 69)	44 (36, 54)
Cape Point	121	0.79	7.6	1.54	27 (18, 39)	27 (19, 39)
Cape Grim	154	0.56	3.4	0.85	12 (9, 15)	14 (11, 17)
South Pole	243	0.75	17.3	5.16	45 (35, 56)	58 (46, 73)

Also listed is the length of the data (in month, M) used for estimation.

shows that in terms of detecting a statistically significant trend, there is no single location that is best for all parameters. Again, the expected change for a variable needs to be compared to the likelihood of detecting a given change in a variable to determine the variables most likely to be useful for the detection of a trend. In the absence of expected or reference trend values, in Figure 2 all sites are being compared on the basis of detecting a 5% per decade trend.

The presentation in this section assumes that one is looking for a trend of a given percent per decade. Often, the expected change is not a percentage change but may be represented as an absolute change, such as an increase of surface ozone of 1 ppb per decade. Table 3 shows some analysis of the surface ozone data also presented in Figure 2, where the data were shown in terms of detecting a 5% per decade trend. Scientifically, it may be more appropriate to look for a fixed change in concentration of surface ozone. Table 3 therefore shows the number of years to detect a 5% per decade trend as well as the number of years to detect a 1 ppb change in surface ozone. Note that the data from Ireland appear to be more appropriate for detecting a 5% per decade trend than the data from Cape Point; however, the data from Cape Point are more appropriate for detecting a trend of 1 ppb per decade than the data from Ireland. Realistically, the uncertainty in the estimate for the number of years for these two sites are not statistically distinguishable. However, this example is only illustrative in nature.

These analyses show the importance of estimating how long it will take to detect trends at different locations and in different environmental parameters. Proper analysis of trend detectability for different locations can help focus existing and future monitoring activities. Choosing sites appropriately can allow scientific questions relevant to the health of the environment to be answered sooner. Efficient monitoring can also save considerably in terms of costs to monitor and delays associated with waiting for reliable trend results.

3. Trend Evaluation With Interventions

Data collection for the detection of trends often requires maintenance of stable instruments in representative locations for decades. However, often unforeseen problems arise which disrupt the measurements in some way that cannot be quantified with any certainty. Instruments may break or may be modified; sites may change location, maintenance and calibration procedures may change, or external events such as volcanic eruptions may occur. A useful statistical model for many of these effects is obtained by assuming that at a specified point in time the data experience a permanent level shift of unknown magnitude [*Box and Tiao*, 1975]. If the data are analyzed in logarithmic form, such a discrete level shift in the logarithms of the original data corresponds to a scale factor change in the original data that could represent, for instance, the change in sensitivity of an instrument. To ignore such events can result in artificial trends in the data that are not representative of the environmental variable being studied. If the magnitude and time of the shift are known, then the data can be adjusted before being analyzed for trends. However, often the time of the level shift is known but not the magnitude. The presence of such a level shift will result in an increase in the variance of a trend estimate that properly accounts for the estimated shift and hence lengthens the time necessary to detect a given trend. The impact of the level shift on trend detection strongly depends on its relative location in time in the data set.

3.1. Basic Statistical Modeling

To investigate, statistically, the effects of intervention level shift, in addition to autocorrelation, on trend estimates, we consider a trend model of the following form:

$$Y_t = \mu + \omega X_t + \delta U_t + N_t, \qquad t = 1, \cdots, T, \tag{5}$$

where ωX_t represents a linear trend beginning at time $t = 0$, as before, and δU_t is a mean level shift term used to account for the possible intervention to the data at the specified time $t = T_0$ ($0 < T_0 < T$); that is,

$$U_t = \begin{cases} 0, & t < T_0 \\ 1, & t \geq T_0. \end{cases}$$

The noise series N_t is modeled as an autoregressive [AR(1)] process, $N_t = \phi N_{t-1} + \varepsilon_t$, as in model (1). We are particularly interested in the uncertainty or variance of the estimate of the trend in model (5), and the effect of occurrence of the intervention mean level shift term (as well as the autocorrelated AR error) on the variance of the estimated trend.

The statistical treatment of level shifts depends to a certain extent on whether the time of the intervention is known and whether any additional information is available regarding the magnitude of the level shift. For the present we consider the case where the time of intervention $t = T_0$ in model (5) is known. In some cases, although the existence of a (possible) intervention within a certain time interval may be expected, the exact time T_0 of the intervention is not known. A variety of statistical work has been done to investigate this problem, including estimation of the intervention time T_0 and testing the null hypothesis of no intervention. Generally, it is found that properties of estimators of the trend and shift coefficients (ω and δ) in model (5) are not affected much when T_0 has to be estimated. The case of additional information about the magnitude of the level shift will be discussed in section 3.3.

It is important to emphasize that model (5) is not universally applicable for all types of interventions. Long-term drifts in instrumentation which are periodically adjusted and volcanic interventions where the effect of the eruption decays with time are examples where model (5) would not be appropriate. However, local changes in sites, maintenance procedures, calibration techniques, and instrumentation may be adequately described by the model considered here. The use of an inappropriate model is likely to result in biased estimates of trends as well as other parameters.

3.2. Trend Estimation

To describe the impact of interventions on trend estimation, let $\hat{\omega}$ denote the GLS estimator of the trend ω in (5), and let $\sigma_{\hat{\omega}}$ = s.d. ($\hat{\omega}$) denote the standard deviation of $\hat{\omega}$. The exact form of $\sigma_{\hat{\omega}}$ is derived in the Appendix, and the expression for the variance ($\sigma^2_{\hat{\omega}}$) of $\hat{\omega}$ is given by (A4). A useful approximation for the standard deviation of the trend estimate is given by

$$\sigma_{\hat{\omega}} \approx \left\{\frac{\sigma_\varepsilon}{(1-\phi)n^{3/2}}\right\} \frac{1}{[1-3\tau(1-\tau)]^{1/2}}$$

$$= \left\{\frac{\sigma_N}{n^{3/2}}\sqrt{\frac{1+\phi}{1-\phi}}\right\} \frac{1}{[1-3\tau(1-\tau)]^{1/2}}, \tag{6}$$

where $\tau = (T_0 - 1)/T$ is the fraction of data before the intervention. This approximation is accurate for smaller values of ϕ but tends to overstate the true value of $\sigma_{\hat{\omega}}$ as ϕ gets larger; also, it becomes more accurate as the number of years of data n gets larger. When there is no intervention, $\tau = 0$, the result in (6) reduces to that given in (2). The effect of a level shift intervention on the standard deviation of the trend estimate ($\sigma_{\hat{\omega}}$) is represented by the factor $1/[1 - 3\tau(1 - \tau)]^{1/2}$, regardless of the value of ϕ. From (6) we see that the uncertainty of a trend estimate is greatest when the intervention occurs halfway through the collection of data ($\tau = 1/2$); in this case, $1/[1 - 3\tau(1 - \tau)]^{1/2} = 2$, so the standard deviation $\sigma_{\hat{\omega}}$ is 2 times the standard deviation obtained in the no-intervention case from (2). When $\tau = 0.25$ or 0.75, the standard deviation $\hat{\sigma}_\omega$ is about 1.5 times the value when no intervention is present.

While the results above show that including a level shift term in the statistical model used will increase the uncertainty of the trend estimate, more serious consequences occur if one does not include a level shift term when a level shift has taken place. If one ignores a potential level shift in the statistical model, the resulting trend estimate will be biased. The exact form of the bias is presented in the Appendix. As an example, in the case of no autocorrelation with $\phi = 0$, it reduces to a bias in the derived trend proportional to the magnitude of the level shift: bias in $\hat{\omega} \approx 6\tau(1 - \tau)/n\ \delta$. In practice, level shifts of moderate magnitude can overwhelm and obscure the true trend being sought, as shown by *Weatherhead et al.* [1997] and *Krzyścin* [1996] in the analysis of UV data. In addition, in practice the estimate of the white noise variance σ^2_ε will be biased upward when the level shift term is not included in the model.

3.3. Extension When There is Additional Information

In some situations, additional (prior) information may be available on the magnitude of the mean level shift δ in model (5). In practice, when the level shift could be caused by instrument-related problems, for example, replacement or recalibration of instrument, this additional information may come from some comparative measurements of the instrument against a standard instrument, or other information available from calibration procedures, a measurement campaign using two instruments for some overlapping time period surrounding the time of intervention. To formalize the nature of information, statistically, we assume that for the magnitude of the level shift (δ) we have a (prior) estimate δ_0 with specified variance (a measure of the precision of prior estimate) σ^2_0. Typically, the prior estimate δ_0 and its variance σ^2_0 will be obtained from independent analysis of the comparative data, for example, simultaneous measurements from two different instruments, mentioned above. We also assume that no other information is available or incorporated into the analysis concerning the other regression parameters μ and ω in model (5).

We denote the ratio $\kappa = \sigma^2_\varepsilon/\sigma^2_0$ as a relative measure of the extent to which we know the magnitude of the level shift. Then, as shown from (A6) in the Appendix, the variance of the corresponding trend estimate $\hat{\omega}$ is given by the expression in (A4) but with h_6 replaced by $h_6 + \kappa$. From this we find that the variance of the trend estimate can be decreased substantially as the amount of additional information increases (i.e., κ increases). To illustrate, consider a series of length $n = 10$ years with the intervention halfway through the time series ($\tau = 0.5$), and moderate autocorrelation of $\phi = 0.5$. When $\kappa = 0$, 1, or 9, the standard deviation of the trend estimate $\sigma_{\hat{\omega}}$ is about 1.8, 1.6, or 1.2, respectively, times that when no intervention is present ($\tau = 0$). Thus when additional information is available corresponding to a value of $\kappa = 9$, the standard deviation of the trend estimate is substantially reduced over the case of no additional information ($\kappa = 0$) and yields variability not too much greater than the case where no intervention is involved, equivalently, where the magnitude of the intervention is known with certainty.

3.4. Trend Detection

Similar to the no-intervention case considered in section 2.3, the number of years n^* of data required to detect a trend of magnitude ω_0 can be determined under the level shift intervention model (5). This model assumes that no additional information on the magnitude of the level shift is available. Using the approximation for the standard deviation of the trend estimate given in (5), a rough but convenient approximation for the number of years for detection is obtained as

$$n^* \approx \left[\frac{3.3\sigma_\varepsilon}{|\omega_0|(1-\phi)[1-3\tau(1-\tau)]^{1/2}} \right]^{2/3}$$
$$= \left[\frac{3.3\sigma_N}{|\omega_0|} \sqrt{\frac{1+\phi}{1-\phi}} \right]^{2/3} \frac{1}{[1-3\tau(1-\tau)]^{1/3}}. \qquad (7)$$

Comparison of this convenient approximation with the approximation (3) for the no-intervention case suggests that the presence of a level shift intervention in the model increases the number of years for trend detection roughly by a factor of $1/[1-3\tau(1-\tau)]^{1/3}$. For example, for the worst case, i.e., $\tau = 1/2$, this approximation gives a factor of 1.59.

Using the exact expression for the standard deviation of the trend estimate given in (A4), the number of years (n^*) required to detect a 5% per decade trend ($\omega_0 = 0.5\%$ per year), for the case of an intervention halfway through the data set ($\tau = 1/2$) are obtained for different values of σ_N and ϕ. By comparing these values with the corresponding values from Table 1 (top), we find that the ratio of the number of years for trend detection under the level shift model to that under the no-intervention model is about 1.55 to 1.59 for smaller values of ϕ ($\phi \le 0.5$), and the ratio becomes slightly smaller for larger values of ϕ. This ratio is not very sensitive to the value of σ_N unless ϕ is large. Thus the presence of an intervention level shift halfway through the data collection causes roughly a 50% increase in the number of years required for detection of trends.

We note that the procedure given in section 2.4 for assessing the uncertainty in the estimated number of years, $\hat{n}^*$, due to the uncertainty in estimating ϕ is applicable when $\hat{n}^*$ is calculated according to (7). This is because the extra factor $1/[1-3\tau(1-\tau)]^{1/3}$ does not depend on ϕ and the large-sample formula for the variance of $\hat{\phi}$ is the same under models (1) and (5).

3.5. Applications

Consider first the characteristics of UV radiation data collected from 14 RB meter stations over the period 1974–1991. In the trend analysis of these UV data by *Weatherhead et al.* [1997], using models similar to those discussed earlier, it was found that the estimates of over 13 of the 14 stations range from about zero to 0.35 with a "typical" value of about 0.2. Also, in units of 100 times the logarithmic data, a "typical" value for the estimate of σ_ε is about 10. The trend estimate using such data can be interpreted roughly as a percentage trend per year. Thus, for example, if we specify that a true trend of magnitude $\omega_0 = 1\%$ per year (as may be expected in the Arctic) is to be detected, then with $\phi = 0.2$ and $\sigma_\varepsilon = 10$, we find from (7) that the required number of years of data for detection of trend of the magnitude $\omega_0 = 1\%$ per year is about $n^* = 11.9$ for the case of no intervention $\tau = 0$, $n^* = 15.7$ for the cases of $\tau = 0.25$ or 0.75, and $n^* = 18.9$ for the case of $\tau = 0.5$. Hence we see from this illustration that the occurrence of level shift in data results in several additional years required for the detection of trend.

We further illustrate the effect of intervention level shifts on trend estimates and their precision using total ozone data from Huancayo, 100 mbar temperature data from Hao, and UV radiation data from Albuquerque. The deseasonalized data for each time series are shown in Figure 3. In each case, a level shift intervention seems highly probable, although the cause or source is not known for all three cases. For each station, trend estimates were obtained for both a no-level shift and a level shift model, which also include a seasonal component, and estimation results are shown in Table 4. We see that in each case there is a substantial difference in trend estimates between the no-level shift and the level shift models.

As discussed in sections 3.1 and 3.2, generally, the inclusion of a level shift term has the effect of increasing the standard deviation of the trend estimate, when other things are equal. However, in practice, if a substantial level shift intervention exists but is not accounted for in the statistical modeling, then the error variance is unduly inflated. Thus inclusion of the level shift term in such prominent cases of intervention will typically result in a decrease in the white noise standard deviation estimate $\hat{\sigma}_\varepsilon$. It may also lead to a reduced estimate $\hat{\phi}$ of the AR(1) coefficient, since some of the low-frequency behavior of the noise that was actually caused by the level shift is then explicitly modeled in the level shift term and hence less low frequency autocorrelation is present in the remaining noise term. As a result of these possible reductions in $\hat{\sigma}_\varepsilon$ and $\hat{\phi}$, in practice the standard deviation of a trend estimate with a level shift term included may not be much larger (or could even be smaller) than that obtained from the model without the level shift term. For the results shown in Table 4, these features tend to be present to some extent.

However, ideally, if the level shift feature had not been present in the data (or could have been "eliminated" in some justifiable way), and the data possessed the same $\hat{\sigma}_\varepsilon$ and $\hat{\phi}$ as obtained from the above level shift model estimation results, then a trend estimate from the no-level shift intervention model would have a substantially smaller standard deviation, as indicated by the results of section 3.2. These "idealized" standard deviation values of the trend estimate under a no-level shift model are shown as $\sigma^*_{\hat{\omega}}$ in column 8 of Table 4. Corresponding values of the estimated number of years $\hat{n}^*$ for detection of trend of magnitude ω_0 are also given in Table 4 for the "idealized" no-intervention and intervention cases. A more complete comparison would be based on interval estimates for n^*, which we omit here as we have illustrated such calculations before.

4. Conclusion

In this paper, the primary statistical considerations in detecting long-term, linear trends were presented. The two major statistical factors governing trend estimation and detection are the autocorrelation and variance of the noise. Results presented show that the number of years of data required to detect a given trend is highly dependent on both of these parameters. Examination of environmental data shows that these two parameters can vary substantially from site to site [e.g., *Zerefos et al.*, 1997] with the result being that one site may require many more years of data to detect a given trend than another site. Additionally, some variables show higher autocorrelation and variability than others, implying that trends are more difficult to detect in some environmental parameters than in others.

Examination of a variety of environmental data sets shows that level shifts are common in long-term monitoring projects as instruments or site locations may change, as well as calibration or maintenance techniques. The work in this paper shows that the occurrence of level shifts in the data can add significantly to the uncertainty in trend estimates and thereby increase the number of years necessary to detect a trend by as much as 50%. However, the effect of level shifts can be minimized in some situations when additional information is avail-

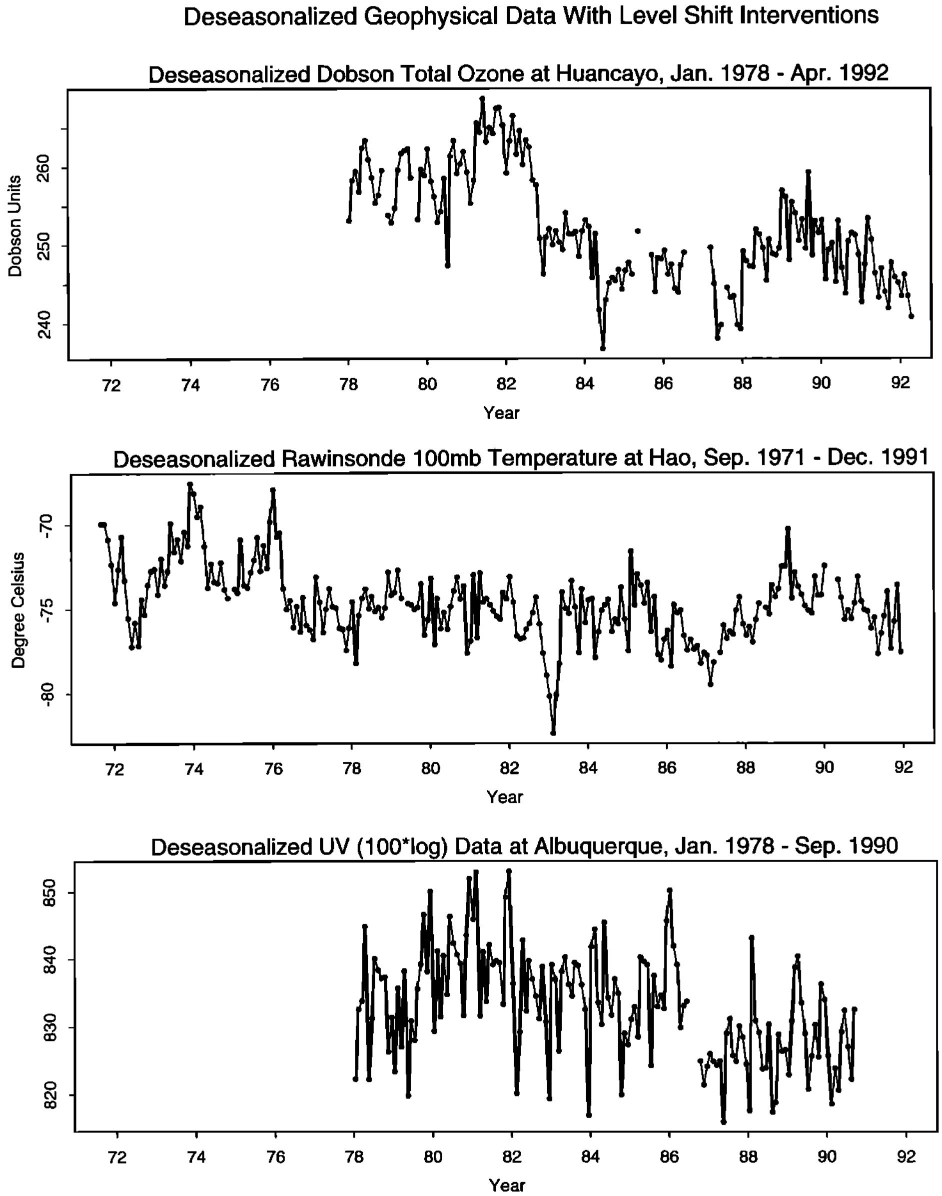

Figure 3. Deseasonalized geophysical data with level shift interventions. The plots show data with identifiable level shifts. Level shifts may indicate an overall change in the parameter measured, a local change not indicative of the parameter over a larger area or a problem in instrumentation.

able in order to estimate the magnitude of the level shift; and the effect can essentially be removed if the magnitude of the level shift is known to a high degree of certainty. Such additional information could come from an overlap or cross calibration in instrumentation, when changes in instruments are made. Estimates are made to show how varying amounts of information mitigate the impact of level shifts on trend estimation. This analysis can be applied to a variety of practical decisions such as to determine the optimal number of months for satellite overlap or calibration routines.

The results presented in this paper are relevant to realistic planning of monitoring sites established for the detection of trends. While there are a variety of other considerations, including the relevance of a site to the scientific question being asked and the logistic maintenance of a monitoring site, the statistical considerations presented in this paper may be useful for determining reasonable expectations for trend detectability at different sites. The applications presented also indicate the great need of additional explanatory data and focused studies to detect environmental change. Small changes of a few percent per decade will take prohibitively long time periods to detect trends, often more than several decades. Further assistance in the detection of trends may be obtained from the use of networks, rather than single monitoring sites. Savings, in terms of number of years to detect a trend, depend on the location and cross correlation of the data from the different sites. To what extent networks and additional explanatory data can help to reduce the number of years for detection is a topic currently under investigation.

Determining which environmental variables and locations are likely to allow for earlier detectable changes will allow current and planned monitoring programs to be more efficient in answering scientific questions. Efficient detection of trends through focused monitoring activities can allow for scientific

Table 4. Illustration of Effect of Intervention Level Shift on Trend Estimate and Its Precision, Based on Geophysical Data From Three Stations

Station	Latitude	$\hat{\phi}$	$\hat{\sigma}_\varepsilon$	$\hat{\delta}$	$\hat{\omega}$	$\hat{\sigma}_{\hat{\omega}}$	$\hat{\sigma}^*_{\hat{\omega}}$
Huancayo (ozone in DU)	12.1°S						
Model without shift		0.715	3.816		−1.068	0.237	
Model with shift		0.527	3.574	−11.260	−0.040	0.228	0.137
(Shift at 11/82; $\tau = 0.34$)				(1.976)		$\hat{n}^* = 12$	$\hat{n}^* = 8$
Hao (temperature in °C)	18.1°S						
Model without shift		0.637	1.563		−0.164	0.046	
Model with shift		0.552	1.513	−3.132	−0.001	0.052	0.036
(Shift at 4/76; $\tau = 0.23$)				(0.705)		$n^* = 29$	$n^* = 23$
Albuquerque (UV in %)	35.1°N						
Model without shift		0.277	6.993		−0.854	0.212	
Model with shift		0.193	6.775	−9.328	0.098	0.320	0.189
(Shift at 9/86; $\tau = 0.69$)				(2.546)		$\hat{n}^* = 18$	$\hat{n}^* = 13$

Estimates and standard deviations in Dobson units (DU) for total ozone, °C for temperature, and % for UV radiation. Notation: $\hat{\phi}$, estimated autocorrelation; $\hat{\sigma}_\varepsilon$ estimated standard deviation of the white noise; $\hat{\delta}$, estimated level shift (with its estimated standard deviation in the parentheses directly underneath it); $\hat{\omega}$, estimated trend; $\hat{\sigma}_{\hat{\omega}}$, estimated standard deviation of the estimated trend; $\hat{\sigma}^*_{\hat{\omega}}$, estimated "idealized" standard deviation of the estimated trend; τ, the fraction of data before the intervention.

questions about changes to the environment to be answered more quickly. Efficient monitoring may be used to show that trends, which may have been expected, have not occurred. Savings from efficient monitoring will include reduced monitoring costs as well as reduction in environmental impact due to early detection of changes.

Appendix

A1. Data Sources

The UV data are from the 14 stations from the original Robertson-Berger UV monitoring network reported on by *Weatherhead et al.* [1997]. The length of the UV data records range from 10 to 18 years. The column ozone data are version 7 from the total ozone mapping spectrometer on the Nimbus 7 satellite. The estimates were based on data from 1979 to 1993. The 14 locations for total ozone were chosen to coincide with the 14 UV monitoring stations. The surface ozone data are from a network collected at the NOAA Climate Monitoring and Diagnostic Laboratory; some stations are analyzed by *Oltmans and Levy* [1994]. The length of the data records range from 4 to 24 years. Data can be obtained from S. Oltmans. The humidity, pressure, and solar radiation data are from the National Renewable Energy Laboratory (NREL) data set which includes some modeled data not expected to affect the results presented here [*NREL*, 1992]. The data are from years 1961 to 1991 inclusive.

A2. Derivation for the Interval Procedure Given in Section 2.4

To derive a 95% confidence interval for n^* when ϕ is estimated by $\hat{\phi}$, we first write explicitly $n^* = n^*(\phi)$. Consequently, when ϕ is estimated by $\hat{\phi}$, n^* is estimated by $\hat{n}^* = n^*(\hat{\phi})$. To construct a 95% interval for n^*, we will use a normal approximation on the log n^* scale. The log scale is used because it helps to improve the normal approximation as well as to maintain the positivity of the resulting (interval) estimate. By the well-known δ method (i.e., first-order Taylor expansion), the variance $\tilde{V}$ of $\log n^*(\hat{\phi})$ is given by

$$\tilde{V} \approx \left(\frac{d \log n^*(\phi)}{d\phi}\right)^2 \times \text{Var}(\hat{\phi}) = \left(\frac{2}{3(1-\phi)}\right)^2 \frac{1-\phi^2}{M}.$$

Since an approximate 95% interval for $\log n^*$ is given by $\log n^*(\hat{\phi}) \pm 2\sqrt{\tilde{V}}$, an approximate 95% interval for n^* is given by $(\hat{n}^* e^{-2\sqrt{\tilde{V}}}, \hat{n}^* e^{2\sqrt{\tilde{V}}})$. The B given in (4) is an estimate of $2\sqrt{\tilde{V}}$ by replacing the unknown ϕ by $\hat{\phi}$. (Strictly speaking, this replacement introduces additional uncertainty. More sophisticated interval estimate based on the so-called variance-stabilizing transformation is available, but the formula is more involved.)

A3. Variance and Bias Calculations Under Models (5) and (1)

With $\mathbf{Y} = (Y_1, \cdots, Y_T)'$ as the $T \times 1$ vector of observations, the trend model (5) with intervention level shift term included may be expressed in matrix form as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{N}, \tag{A1}$$

where $\mathbf{X}$ is a $T \times 3$ matrix consisting of the constant, trend, and intervention terms, $\boldsymbol{\beta} = (\mu, \omega, \delta)'$ represents the vector of unknown regression coefficients to be estimated, and $\mathbf{N} = (N_1, \cdots, N_T)'$ is the $T \times 1$ vector of noise terms. For the model above with AR(1) noise process $\{N_t\}$, let $\boldsymbol{\varepsilon} = (\sqrt{1-\phi^2}N_1, \varepsilon_2, \cdots, \varepsilon_T)'$, which has $\text{Cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2\mathbf{I}$. From the AR(1) equation $N_t - \phi N_{t-1} = \varepsilon_t$, it follows that the noise vector $\mathbf{N}$ satisfies $\mathbf{P}'\mathbf{N} = \boldsymbol{\varepsilon}$, so that $\mathbf{N} = \mathbf{P}'^{-1}\boldsymbol{\varepsilon}$, where the matrix $\mathbf{P}'$ is $T \times T$ with (1, 1)-element equal to $\sqrt{1-\phi^2}$, the remaining diagonal elements equal to 1, the $(i, i-1)$-elements equal to $-\phi$, and zero elements otherwise. Hence the covariance matrix of $\mathbf{N}$ has the form $\text{Cov}(\mathbf{N}) = \text{Cov}(\mathbf{P}'^{-1}\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2\mathbf{P}'^{-1}\mathbf{P}^{-1} \equiv \sigma_\varepsilon^2\mathbf{V}$, with $\mathbf{V} = \mathbf{P}'^{-1}\mathbf{P}^{-1}$ or $\mathbf{V}^{-1} = \mathbf{P}\mathbf{P}'$. Consider the transformed equation

$$\mathbf{Y}^* = \mathbf{P}'\mathbf{Y} = \mathbf{P}'\mathbf{X}\boldsymbol{\beta} + \mathbf{P}'\mathbf{N} = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{A2}$$

where $\mathbf{X}^* = \mathbf{P}'\mathbf{X}$, and $\text{Cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2\mathbf{I}$. The generalized least squares (GLS) estimator of $\boldsymbol{\beta}$ in model (A1) is then the ordinary least squares estimator in the transformed model, $\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{Y}^*$, with covariance matrix

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \sigma_\varepsilon^2(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}. \tag{A3}$$

After some algebra, we have

$$\mathbf{X}^{*\prime}\mathbf{X}^* = \begin{bmatrix} h_1 & h_2 & h_4 \\ h_2 & h_3 & h_5 \\ h_4 & h_5 & h_6 \end{bmatrix},$$

where

$$h_1 = (T-1)(1-\phi)^2 + (1-\phi^2),$$

$$h_2 = \frac{1-\phi}{12}\left[\tfrac{1}{2}T(T-1)(1-\phi) + T + \phi\right],$$

$$h_3 = \tfrac{1}{144}\left[\tfrac{1}{6}T(T+1)(2T+1)(1-\phi)^2 + T^2\phi(1-\phi) + T\phi - \phi^2\right],$$

$$h_4 = (T-T_0)(1-\phi)^2 + (1-\phi),$$

$$h_5 = \frac{(T-T_0)(1-\phi)}{24}\left[(T+T_0)(1-\phi) + 1 + \phi\right] + \tfrac{1}{12}\left[T_0 - (T_0-1)\phi\right],$$

$$h_6 = (T-T_0)(1-\phi)^2 + 1.$$

Therefore using matrix algebra, Var $(\hat{\omega})$ is obtained as σ_ε^2 times the (2, 2)-element of the inverse of the matrix $\mathbf{X}^{*\prime}\mathbf{X}^*$, which yields the expression

$$\mathrm{Var}(\hat{\omega}) = \sigma_\varepsilon^2 \frac{h_6(h_1h_6 - h_4^2)}{(h_1h_6 - h_4^2)(h_3h_6 - h_5^2) - (h_2h_6 - h_4h_5)^2} \equiv \sigma_\varepsilon^2 h_*^2(\phi, T, T_0). \tag{A4}$$

In the special case of no autocorrelation when $\phi = 0$, expression (A4) simplifies to

$$\mathrm{Var}(\hat{\omega}) = \sigma_\varepsilon^2 \frac{12^3}{2T(T+1)(2T+1) - 3(T-T_0+1)(T+T_0)^2 - 3T_0^2(T_0-1)}.$$

A close approximation to the variance for this special case is Var $(\hat{\omega}) \approx \sigma_\varepsilon^2 12^3/\{T^3[1 - 3\tau(1-\tau)]\} = \sigma_\varepsilon^2/\{n^3[1 - 3\tau(1-\tau)]\}$, where $n = T/12$ is the number of years of data, and $\tau = (T_0 - 1)/T$ is the fraction of preintervention data. In a similar way, for a general value of ϕ, a useful approximation for (A4) can be obtained as Var $(\hat{\omega}) \approx \sigma_\varepsilon^2/\{(1-\phi)^2 n^3[1 - 3\tau(1-\tau)]\}$, but this approximation is not so accurate for larger values of ϕ.

The result for model (1) can be obtained as a special case, since (1) results from omitting the level shift term δU_t from (5). Let $\mathbf{X} = [\mathbf{X}_1, \mathbf{U}]$ where $\mathbf{X}_1$ represents the $T \times 2$ matrix of regressors for model (1), and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \delta)'$ where $\boldsymbol{\beta}_1 = (\mu, \omega)'$. Then the GLS estimator of $\boldsymbol{\beta}_1$ under model (1) is $\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1^{*\prime}\mathbf{X}_1^*)^{-1}\mathbf{X}_1^{*\prime}\mathbf{Y}^*$, with covariance matrix

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}_1) = \sigma_\varepsilon^2(\mathbf{X}_1^{*\prime}\mathbf{X}_1^*)^{-1} \equiv \sigma_\varepsilon^2\begin{bmatrix} h_1 & h_2 \\ h_2 & h_3 \end{bmatrix}^{-1},$$

where $\mathbf{X}_1^* = \mathbf{P}'\mathbf{X}_1$. So, for model (1) it follows immediately that Var $(\hat{\omega})$ is given by σ_ε^2 times the (2, 2)-element of the inverse of the matrix $\mathbf{X}_1^{*\prime}\mathbf{X}_1^*$,

$$\mathrm{Var}(\hat{\omega}) = \sigma_\varepsilon^2 \frac{h_1}{h_1h_3 - h_2^2} \equiv \sigma_\varepsilon^2 h^2(\phi, T). \tag{A5}$$

For the special case of no autocorrelation when $\phi = 0$, the expression simplifies to Var $(\hat{\omega}) = \sigma_\varepsilon^2 12^3/\{T(T^2-1)\}$. A useful, simple approximation for the variance expression in (A5) is Var $(\hat{\omega}) \approx \sigma_\varepsilon^2 12^3/\{(1-\phi)^2 T(T^2-1)\}$, which is quite accurate when $|\phi|$ is not close to 1.

Next, we consider the bias of the GLS estimator $\hat{\boldsymbol{\beta}}_1$ obtained by assuming model (1), without intervention term, when the true model is (5) with a nonzero level shift intervention term. Under model (5), $E[\mathbf{Y}^*] = \mathbf{X}_1^*\boldsymbol{\beta}_1 + \mathbf{U}^*\delta$, so the expected value of $\hat{\boldsymbol{\beta}}_1$ is

$$E[\hat{\boldsymbol{\beta}}_1] = (\mathbf{X}_1^{*\prime}\mathbf{X}_1^*)^{-1}\mathbf{X}_1^{*\prime}E[\mathbf{Y}^*] = \boldsymbol{\beta}_1 + (\mathbf{X}_1^{*\prime}\mathbf{X}_1^*)^{-1}\mathbf{X}_1^{*\prime}\mathbf{U}^*\delta \equiv \boldsymbol{\beta}_1 + \begin{bmatrix} h_1 & h_2 \\ h_2 & h_3 \end{bmatrix}^{-1}\begin{bmatrix} h_4 \\ h_5 \end{bmatrix}\delta.$$

Hence we see from this that the expected value of the trend estimate obtained from model (1), when the true model is (5) with a nonzero level shift, is

$$E[\hat{\omega}] = \omega + \frac{h_1h_5 - h_2h_4}{h_1h_3 - h_2^2}\delta.$$

In the case of no autocorrelation with $\phi = 0$, this expression reduces to

$$E[\hat{\omega}] = \omega + \frac{6(T_0-1)(T-T_0+1)}{T(T^2-1)/12}\delta \approx \omega + \frac{6\tau(1-\tau)}{n}\delta.$$

Now we consider the situation of section 3.2 where it is assumed that for δ in model (5) we have additional information in the form of a (prior) estimate δ_0 with specified variance σ_0^2. Set $\boldsymbol{\beta}_0 = (0, 0, \delta_0)'$ with Σ_0^- denoting the 3×3 matrix whose elements are zero except for the (3, 3)-element which is equal to $1/\sigma_0^2$. Then the efficient estimate of $\boldsymbol{\beta}$ in (A1), which incorporates the above prior estimate or additional information, is given by $\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*\prime}\mathbf{X}^* + \sigma_\varepsilon^2\Sigma_0^-)^{-1}(\mathbf{X}^{*\prime}\mathbf{Y}^* + \sigma_\varepsilon^2\Sigma_0^-\boldsymbol{\beta}_0)$ with covariance matrix

$$\Sigma_\beta \equiv \mathrm{Cov}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sigma_\varepsilon^2(\mathbf{X}^{*\prime}\mathbf{X}^* + \sigma_\varepsilon^2\Sigma_0^-)^{-1}. \tag{A6}$$

Notation

- Y_t time series of environmental data.
- S_t seasonal mean function.
- X_t linear trend function.
- N_t AR(1) noise.
- ε_t white noise.
- μ mean of data.
- ω_0 trend to be detected.
- $\hat{\omega}$ trend estimate.
- $\sigma_{\hat{\omega}}$ uncertainty of trend estimate.
- σ_N standard deviation of the AR(1) noise.
- $\hat{\sigma}_N$ estimate of σ_N.
- σ_ε standard deviation of the white noise.
- $\hat{\sigma}_\varepsilon$ estimate of σ_ε.
- ϕ autocorrelation of the AR(1) noise.
- $\hat{\phi}$ estimate of ϕ.
- T_0 time of intervention or level shift.
- T length of data set in months.
- n length of data set in years.
- n^* years to detect a trend.
- $\hat{n}^*$ estimate of n^*.

Acknowledgments. This research was supported by the U.S. Department of Energy under grant DE-FG03-94ER61857. We are grateful for their support. The assistance of Gene Maxwell in providing data and advice is also appreciated.

References

Box, G. E. P., and G. C. Tiao, Intervention analysis with applications to economic and environmental problems, *J. Am. Stat. Assoc.*, *70*, 70–79, 1975.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed., Prentice-Hall, Englewood Cliffs, N. J., 1994.

Krzyścin, J. W., UV controlling factors and trends derived from the ground-based measurements taken at Belsk, Poland, 1976–1994, *J. Geophys. Res.*, *101*, 16,797–16,805, 1996.

Miller, A. J., R. M. Nagatani, G. C. Tiao, X. F. Niu, G. C. Reinsel, D. Wuebbles, and K. Grant, Comparisons of observed ozone and temperature trends in the lower stratosphere, *Geophys. Res. Lett.*, *19*, 929–932, 1992.

National Renewable Energy Laboratory (NREL), *National Solar Radiation Data Base (1961–1990) User's Manual*, version 1.0, Golden, Colo., 1992.

Oltmans, S. J., and H. Levy II, Surface ozone measurements from a global network, *Atmos. Environ.*, *28*, 9–22, 1994.

Reinsel, G. C., G. C. Tiao, D. J. Wuebbles, J. B. Kerr, A. J. Miller, R. M. Nagatani, L. Bishop, and L. H. Ying, Seasonal trend analysis of published ground-based and TOMS total ozone data through 1991, *J. Geophys. Res.*, *99*, 5449–5464, 1994.

Scotto, J., G. Cotton, F. Urbach, D. Berger, and T. Fears, Biologically effective ultraviolet radiation: Surface measurements in the United States, 1974 to 1985, *Science*, *239*, 762–764, 1988.

Stolarski, R. S., P. Bloomfield, and R. D. McPeters, Total ozone trends deduced from Nimbus 7 TOMS data, *Geophys. Res. Lett.*, *18*, 1015–1018, 1991.

Stolarski, R., R. Bojkov, L. Bishop, C. Zerefos, J. Staehelin, and J. Zawodny, Measured trends in stratospheric ozone, *Science*, *256*, 342–349, 1992.

Tiao, G. C., G. C. Reinsel, D. Xu, J. H. Pedrick, X. Zhu, A. J. Miller, J. J. DeLuisi, C. L. Mateer, and D. J. Wuebbles, Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation, *J. Geophys. Res.*, *95*, 20,507–20,517, 1990.

Weatherhead, E. C., G. C. Tiao, G. C. Reinsel, J. E. Frederick, J. J. DeLuisi, D. Choi, and W.-K. Tam, Analysis of long-term behavior of ultraviolet radiation measured by Robertson-Berger meters at 14 sites in the United States, *J. Geophys. Res.*, *102*, 8737–8754, 1997.

Zerefos, C. S., D. S. Balis, A. F. Bais, D. Gillotay, P. C. Simon, B. Mayer, and G. Seckmeyer, Variability of UV-B at four stations in Europe, *Geophys. Res. Lett.*, *24*, 1363–1366, 1997.

W.-K. Cheang and G. C. Reinsel, Department of Statistics, University of Wisconsin, Madison, WI 53706.

D. Choi and X.-L. Meng, Department of Statistics, University of Chicago, Chicago, IL 60637.

J. DeLuisi, NOAA ARL, 325 Broadway, Boulder, CO 80303.

J. E. Frederick, Department of Geophysical Sciences, University of Chicago, Chicago, IL 60637.

T. Keller, National Center for Atmospheric Research, Boulder, CO 80307.

J. B. Kerr, Atmospheric Environment Service, 4905 Dufferin Street, Downsview, Ontario, Canada.

A. J. Miller, NOAA NWS, 5200 Auth Road, Washington, D. C. 20233.

S. J. Oltmans, NOAA CMDL, 325 Broadway, Boulder, CO 80303.

G. C. Tiao, Graduate School of Business, University of Chicago, Chicago, IL 60637.

E. C. Weatherhead, Cooperative Institute for Research in the Environmental Sciences, University of Colorado, 325 Broadway, Boulder, CO 80309. (e-mail: betsy@srrb.noaa.gov)

D. J. Wuebbles, Department of Atmospheric Sciences, University of Illinois, Champaign, IL 61801.

(Received November 10, 1997; revised March 2, 1998; accepted March 4, 1998.)