

## A Comparison of Climate Signal Trend Detection Uncertainty Analysis Methods

NIPA PHOJANAMONGKOLKIJ

*Systems Engineering and Engineering Methods Branch, NASA Langley Research Center, Hampton, Virginia*

SEIJI KATO, BRUCE A. WIELICKI, PATRICK C. TAYLOR, AND MARTIN G. MLYNCZAK

*Climate Science Branch, NASA Langley Research Center, Hampton, Virginia*

(Manuscript received 9 July 2013, in final form 12 December 2013)

### ABSTRACT

Two climate signal trend analysis methods are the focus of this paper. The uncertainty of trend estimate from these two methods is investigated using Monte Carlo simulation. Several theoretically and randomly generated series of white noise, first-order autoregressive and second-order autoregressive, are explored. The choice of method that is most appropriate for the time series of interest depends upon the autocorrelation structure of the series. If the structure has its autocorrelation coefficients decreased with increasing lags (i.e., an exponential decay pattern), then the method of Weatherhead et al. is adequate. If the structure exhibits a decreasing sinusoid pattern of coefficient with lags (or a damped sinusoid pattern) or a mixture of both exponential decay and damped sinusoid patterns, then the method of Leroy et al. is recommended. The two methods are then applied to the time series of monthly and globally averaged top-of-the-atmosphere (TOA) irradiances for the reflected solar shortwave and emitted longwave regions, using radiance observations made by Clouds and the Earth's Radiant Energy System (CERES) instruments during March 2000 through June 2011. Examination of the autocorrelation structures indicates that the reflected shortwave region has an exponential decay pattern, while the longwave region has a mixture of exponential decay and damped sinusoid patterns. Therefore, it is recommended that the method of Weatherhead et al. is used for the series of reflected shortwave irradiances and that the method of Leroy et al. is used for the series of emitted longwave irradiances.

### 1. Introduction

Detection of long-term trends in various types of atmospheric data has been studied extensively in the past 30 years (e.g., Hill et al. 1977; Reinsel et al. 1989; Tiao et al. 1990; Weatherhead et al. 1998; Von Storch and Zwiers 1999; Leroy et al. 2008a,b). There are many factors that affect the detection, including the magnitude of the trend to be detected, the temporal length of available observations, the magnitude of the variability and autocorrelation of the noise in the data, and uncertainty in the observations. During the past decade, two trend detection methods have been widely used: the method of Weatherhead et al. (1998, hereafter W98) and

the method of Leroy et al. (2008b, hereafter L08). It was not clear if the uncertainty of climate trends and the time to detect them are sensitive to the method used and under what circumstances that one method is preferred over the other. That was the motivation for this study. The W98 method is detailed in section 1a. Section 1b details the L08 method. Comparison of both methods is given in section 1c.

#### a. The W98 method

Let  $Y_t$  be the measurement of interest. The trend model assumes linearity, as expressed in  $Y_t = \alpha + S_t + \omega X_t + N_t$ , where  $\alpha$  is a constant term,  $S_t$  is the seasonal component,  $X_t = t/12$  is the linear trend function,  $\omega$  is the magnitude of the trend per year,  $N_t$  is the unexplained noise of the measurement data, and  $t = 1, 2, 3, \dots, T$  months. The noise  $N_t$  can be obtained from the deseasonalized and detrended time series monthly data. A deseasonalized series is obtained by subtracting the monthly-mean value from each month. For example, the average of all January months is subtracted from each

 Denotes Open Access content.

Corresponding author address: Nipa Phojanamongkolkij, NASA Langley Research Center, MS 131, Hampton, VA 23681-2199.  
E-mail: nipa.phojanamongkolkij@nasa.gov

DOI: 10.1175/JCLI-D-13-00400.1

individual January month, with the same process used for all months in the year. This process removes the bulk of Earth's very large and repeatable seasonal cycle from the climate record. Then the trend is removed from the deseasonalized series, generating the  $N_t$  deseasonalized and detrended time series.

The W98 method assumes the noise  $N_t$  follows a first-order autoregressive model (AR1). The uncertainty of trend estimate given by the W98 is

$$[\text{Var}(\hat{\omega})]^{1/2} \approx \left[ \frac{\text{Var}(N_t)}{n^3} \left( \frac{1 + \rho_1}{1 - \rho_1} \right) \right]^{1/2}. \quad (1)$$

The estimate of the trend per year is  $\hat{\omega}$ ,  $\text{Var}(\hat{\omega})$  is the variance of the trend estimate,  $n = T/12$  is the number of years of time series data,  $\text{Var}(N_t)$  is the variance of the deseasonalized and detrended  $N_t$  monthly time series, and  $\rho_1 = \text{Corr}(N_{t+1}, N_t)$  is the lag-1 autocorrelation coefficient.

### b. The L08 method

The trend model for the L08 is also linear, as expressed in  $Y_i = \alpha + mt_i + N_i$ , where  $Y_i$  is the measurement,  $\alpha$  is a constant term,  $t_i$  is the linear trend variable,  $m$  is the magnitude of the trend per month,  $N_i$  is the unexplained noise of the measurement data, and  $i = 1, 2, 3, \dots, T$  months. The uncertainty of trend estimate given by the L08 is

$$[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{(\Delta t)^3} \text{Var}(N_i) \left( \sum_{\mu=-\infty}^{\infty} \rho_{\mu} \right) \right]^{1/2}. \quad (2)$$

Term  $\Delta t = Tdt$  is the length of time series, and  $dt$  is the interval between observations. In this study  $dt = 1$ , and therefore  $\Delta t$  reduces to  $T$ .

The L08 takes theoretical autocorrelation at all lags (i.e., from  $-\infty$  to  $+\infty$ ) into consideration to obtain the uncertainty of the trend estimate [Eq. (2)]. However, we do not know the theoretical autocorrelation  $\rho_{\mu}$ , but can estimate it by the sample autocorrelation  $\hat{\rho}_{\mu}$  from the available time series data. Hassani (2010) has shown that  $\sum_{\mu=1}^{T-1} \hat{\rho}_{\mu} = -1/2$ . We also know the properties of  $\hat{\rho}_{-\mu} = \hat{\rho}_{\mu}$  and  $\hat{\rho}_0 = 1$ . Therefore, it is not possible to estimate  $\sum_{\mu=-\infty}^{\infty} \rho_{\mu}$  with  $\sum_{\mu=-(T-1)}^{T-1} \hat{\rho}_{\mu}$  in practice because  $\sum_{\mu=-(T-1)}^{T-1} \hat{\rho}_{\mu} = 0$ . Consequently, a criterion is needed for which sample autocorrelation  $\hat{\rho}_{\mu}$  to consider.

### c. Method comparison

The W98 has its uncertainty for trend estimate per year [Eq. (1)], while the L08 has its uncertainty in a per month basis [Eq. (2)]. To make the W98 into the uncertainty for trend estimate per month and assume the trend is reasonably linear with time, we take the

variance property of  $\text{Var}(\hat{m}) = \text{Var}(\hat{\omega}/12) = \text{Var}(\hat{\omega})/12^2$ , and therefore Eq. (1) becomes

$$\begin{aligned} \text{Var}(\hat{m}) &\approx \frac{1}{12^2} \frac{\text{Var}(N_t)}{(T/12)^3} \left( \frac{1 + \rho_1}{1 - \rho_1} \right) \\ &\approx \frac{12}{T^3} \text{Var}(N_t) \left( \frac{1 + \rho_1}{1 - \rho_1} \right) \quad \text{and} \end{aligned}$$

$$[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{T^3} \text{Var}(N_t) \left( \frac{1 + \rho_1}{1 - \rho_1} \right) \right]^{1/2}. \quad (3)$$

The obvious difference in the two uncertainties [Eqs. (2) and (3)] is the last term on the right-hand side of the equations, which indicates how the correlation structure is handled by the two methods. We define decorrelation time  $\tau$  as the time needed for data to be uncorrelated (Von Storch and Zwiers 1999). The decorrelation time for W98 and L08 was set as  $\tau_W = (1 + \rho_1)/(1 - \rho_1)$  and  $\tau_L = \sum_{\mu=-\infty}^{\infty} \rho_{\mu}$ , respectively. Therefore, we focus on comparing terms  $\text{Var}(N)\tau_W$  and  $\text{Var}(N)\tau_L$  when comparing the two methods.

The W98 assumes the noise  $N_t$  follows the AR1 model, which requires only one estimated model parameter (i.e., the first lag sample autocorrelation coefficient  $\hat{\rho}_1$ ). As such, it is simple to use. However, any departure of  $N_t$  from the AR1 model will likely cause the W98 to perform poorly. The L08 does not assume any autoregressive model or any correlation structure for the  $N_t$  series. It takes into account theoretical autocorrelation coefficients at all lags in computing the uncertainty of trend estimate. It is therefore theoretically versatile. However, all theoretical coefficients are not known in practice, and rather are estimated by their corresponding sample autocorrelation coefficients. Mathematically it is not possible to include sample coefficients at all lags because  $\sum_{\mu=-(T-1)}^{T-1} \hat{\rho}_{\mu} = 0$ . Therefore, it is essential to have a criterion on which sample coefficients to include in the estimation of the L08's uncertainty. Table 1 summarizes the equations for the uncertainty of trend estimate and the advantages and disadvantages of both methods.

The organization of the paper is as follows: Section 2 describes and compares the two methods using the Monte Carlo simulation with theoretically and randomly constructed time series. The findings from this section are applied to the reflected shortwave and emitted longwave irradiances at the top of the atmosphere (TOA), as estimated during March 2000 through June 2011 from observed Clouds and the Earth's Radiant Energy System (CERES) instruments on the National Aeronautics and Space Administration (NASA) Terra satellite. This application is detailed in section 3. Section 4 concludes the study.

TABLE 1. Method comparison summary.

Methods	W98	L08
Uncertainty of trend estimate $\hat{m}$ (per month)	From Eq. (3), $[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{T^3} \text{Var}(N_t) \left( \frac{1 + \rho_1}{1 - \rho_1} \right) \right]^{1/2}$ or $[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{T^3} \text{Var}(N_t) \tau_W \right]^{1/2}$	From Eq. (2), $[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{(\Delta t)^3} \text{Var}(N_i) \left( \sum_{\mu=-\infty}^{\infty} \rho_{\mu} \right) \right]^{1/2}$ or $[\text{Var}(\hat{m})]^{1/2} \approx \left[ \frac{12}{T^3} \text{Var}(N_t) \tau_L \right]^{1/2}$
Advantage	The $N_t$ time series simply follows the AR1 process, with one estimated model parameter (i.e., the first lag sample autocorrelation $\hat{\rho}_1$ ) required.	The $N_t$ time series follows any process, and the correlation structure of the series can be adequately captured.
Disadvantage	Uncertainty of trend estimate depends on how adequate the AR1 process captures the correlation structure of the series.	Uncertainty of trend estimate relies on a criterion to determine which sample autocorrelation coefficients to include in the equation. No established criterion is available.

**2. Monte Carlo simulation study**

In section 2a, the Monte Carlo simulation settings are described. Simulation results are given in section 2b.

*a. Simulation settings*

We compare the W98 and L08 methods with the theoretically and randomly generated first- and second-order autoregressive (i.e., AR1 and AR2) time series. The AR1 series follows  $N_t = \phi_1 N_{t-1} + \varepsilon_t$ , and  $\rho_{\mu} = \phi_1^{\mu}$ , where  $N_t$  is a randomly generated observation at time  $t$ ,  $\phi_1$  is the model parameter, and  $\varepsilon_t$  is serially uncorrelated normally distributed random data with a mean of zero and standard deviation of one (or white noise) at time  $t$ , or  $\varepsilon_t \sim \text{Norm}(0, 1)$ . The  $N_t$  is a deseasonalized and detrended time series as described in section 1. Note that the white noise series can be thought of as AR1 with  $\phi_1 = 0$ . For an AR1 series to be stationary, it is required that  $|\phi_1| < 1$  (Montgomery et al. 2008).

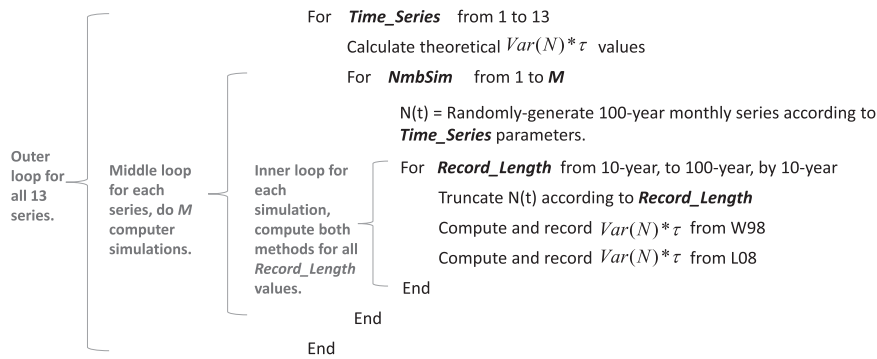
The AR2 series follows  $N_t = \phi_1 N_{t-1} + \phi_2 N_{t-2} + \varepsilon_t$ , and  $\rho_{\mu} = \phi_1 \rho_{\mu-1} + \phi_2 \rho_{\mu-2}$ , where  $\phi_1$  and  $\phi_2$  are parameters

of the model. For an AR2 series to be stationary, we consider two roots  $m_1$  and  $m_2$  of the polynomial  $m^2 - \phi_1 m - \phi_2 = 0$ . There are three possible cases for the two roots. In case 1, if both roots are real, distinct, and  $|m_1|$  and  $|m_2| < 1$ , then the autocorrelation function (ACF) exhibits a mixture of two exponential decay patterns. In case 2, if both roots are complex conjugates in the form of  $a \pm ib$  and satisfy  $\sqrt{a^2 + b^2} < 1$ , then the ACF has a damped sinusoid pattern. In case 3, if there is one real root with  $m_1 = m_2$ , then the ACF shows a single exponential decay pattern. We study all three cases of the AR2 series. We limit our simulation study to AR2 for simplicity because it already captures phenomena that are likely to complicate the W98.

Table 2 outlines the time series data that are used with the W98 and L08 methods in the first three columns. The sensitivity of both methods to record length (for 10, 20, . . . , 100 years) is included in this study. The exploration matrix is a 13 time series by two methods by 10 record lengths (13 × 2 × 10) matrix. The column “series number” will be used to facilitate all output figures.

TABLE 2. Exploration matrix for the W98 and L08 methods. The record lengths for time series data are 10, 20, . . . , 100 yr.

Time series data	ACF pattern	Model parameters	Series no.
White noise	None	$N_t = \varepsilon_t$ , where $\phi_1$ and $\phi_2 = 0$	AR1_0
AR1	Single exponential decay	$N_t = 0.3N_{t-1} + \varepsilon_t$ , where $\phi_1 = 0.3$ and $\phi_2 = 0$	AR1_1
		$N_t = 0.6N_{t-1} + \varepsilon_t$ , where $\phi_1 = 0.6$ and $\phi_2 = 0$	AR1_2
		$N_t = 0.9N_{t-1} + \varepsilon_t$ , where $\phi_1 = 0.9$ and $\phi_2 = 0$	AR1_3
AR2	Mixture of two exponential decay patterns	Case 1: $N_t = 0.3N_{t-1} + 0.1N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.3$ and $\phi_2 = 0.1$	AR2_1a
		Case 1: $N_t = 0.6N_{t-1} + 0.1N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.6$ and $\phi_2 = 0.1$	AR2_1b
		Case 1: $N_t = 0.9N_{t-1} - 0.1N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.9$ and $\phi_2 = -0.1$	AR2_1c
	Damped sinusoid	Case 2: $N_t = 0.3N_{t-1} - 0.5N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.3$ and $\phi_2 = -0.5$	AR2_2a
		Case 2: $N_t = 0.6N_{t-1} - 0.5N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.6$ and $\phi_2 = -0.5$	AR2_2b
		Case 2: $N_t = 0.9N_{t-1} - 0.5N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.9$ and $\phi_2 = -0.5$	AR2_2c
	Single exponential decay	Case 3: $N_t = 0.3N_{t-1} - 0.0225N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.3$ and $\phi_2 = -0.0225$	AR2_3a
		Case 3: $N_t = 0.6N_{t-1} - 0.09N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.6$ and $\phi_2 = -0.09$	AR2_3b
		Case 3: $N_t = 0.9N_{t-1} - 0.2025N_{t-2} + \varepsilon_t$ , where $\phi_1 = 0.9$ and $\phi_2 = -0.2025$	AR2_3c



where  $N(t)$  = Randomly-generated 100-year monthly series

FIG. 1. Computer simulation framework.

Figure 1 shows in detail how the computer simulation is done. The outer loop is repeated 13 times, one for each time series. The middle loop is repeated  $M$  times, one for each simulation. The determination of the number of simulations  $M$  will be given later in this section. For each simulation, the inner loop computes the  $\text{Var}(N)\tau$  values by the two methods and is repeated for 10 different record lengths of interest. Also produced for each time series is the theoretical  $\text{Var}(N)\tau$  value, which does not depend on record length. We choose to compare the  $\text{Var}(N)\tau$  terms rather than the uncertainties of trend estimate because the uncertainties are functions of both record length  $T$  and  $\text{Var}(N)\tau$  term [Eqs. (2) and (3)]. Given significantly large  $T$ , the two methods could provide similar and very small uncertainties, but the similarity is primarily attributable to the effect of large  $T$ . Therefore, comparing the  $\text{Var}(N)\tau$  with respect to the theoretical value using two methods is adequate to get an insight into their performances independently from the record length  $T$ .

The theoretical value for AR1 from Eq. (3) is

$$\text{Var}(N_t)\tau = \text{Var}(N_t) \left( \frac{1 + \rho_1}{1 - \rho_1} \right). \quad (4)$$

According to W98,  $\text{Var}(N_t) = \text{Var}(\varepsilon_t)/(1 - \rho_1^2)$ , where  $\text{Var}(\varepsilon_t)$  is the variance of serially uncorrelated normally distributed random data, which in this study is one. Given that  $1 - \rho_1^2 = (1 - \rho_1)(1 + \rho_1)$ , Eq. (4) becomes

$$\text{Var}(N_t)\tau = \frac{\text{Var}(\varepsilon_t)}{(1 - \rho_1)^2}. \quad (5)$$

The theoretical  $\text{Var}(N)\tau$  value for the AR2 is [from Eq. (A12) in the appendix]

$$\text{Var}(N_t)\tau = \frac{\text{Var}(\varepsilon_t)}{(1 - \phi_1 - \phi_2)^2}. \quad (6)$$

Parameters  $\phi_1$  of the AR1 model in Table 2 are chosen to be 0.0, 0.3, 0.6, and 0.9, representing white noise, low-, medium-, and high-correlated series. For each AR1 series, we first randomly generate a data point sampled from a  $\text{Norm}(0, 1)$  distribution. This first sample represents a data at time  $t = 0$ , or  $N_0 = \varepsilon_0$ . The next data point  $N_1$  is obtained from  $\phi_1 N_0 + \varepsilon_1$ . The process is repeated until 1200 data points (from  $N_1$  to  $N_{1200}$ ) are obtained, representing a 100-yr monthly series. The ACFs of all chosen 100-yr monthly AR1 models are given in Fig. 2. Sample autocorrelation coefficients up to  $10 \log_{10}(T)$  lags, where  $T = 1200$ , are computed and shown in the ACFs (Venables and Ripley 2002). Any autocorrelations that are outside the dashed lines are considered statistically significant within a 95% confidence level. For the white noise series (Fig. 2a), only lag 0 is significant, which is the property of the serially uncorrelated series. Figures 2b to 2d give ACFs exhibiting exponentially decay patterns. As parameter  $\phi_1$  increases, its ACF pattern decays slower as lags increase.

Parameters  $\phi_1$  and  $\phi_2$  of the AR2 model are chosen to represent the three cases of the stationary AR2 series. For each AR2 series, we first randomly generate two data points  $N_{-1}$  and  $N_0$  sampled from a  $\text{Norm}(0, 1)$  distribution (i.e.,  $\varepsilon_{-1}$  and  $\varepsilon_0$ ). The next data point  $N_1$  is obtained from  $\phi_1 N_0 + \phi_2 N_{-1} + \varepsilon_1$ . The process is repeated until 1200 data points (from  $N_1$  to  $N_{1200}$ ) are obtained, representing a 100-yr monthly series. The ACF of the first case has a mixture of two exponential decay patterns (Figs. 3a–c). The ACF of the second case exhibits a damped sinusoid pattern (Figs. 3d–f). The ACF of the third case has a single exponential decay pattern (Figs. 3g–i).

We determine the number of simulations ( $M$  in Fig. 1) needed for all chosen time series using the variance reduction technique (Ross 2006). The average  $\text{Var}(N)\tau$  of the simulation solutions has to be an unbiased estimator of its corresponding truth. This average value changes as

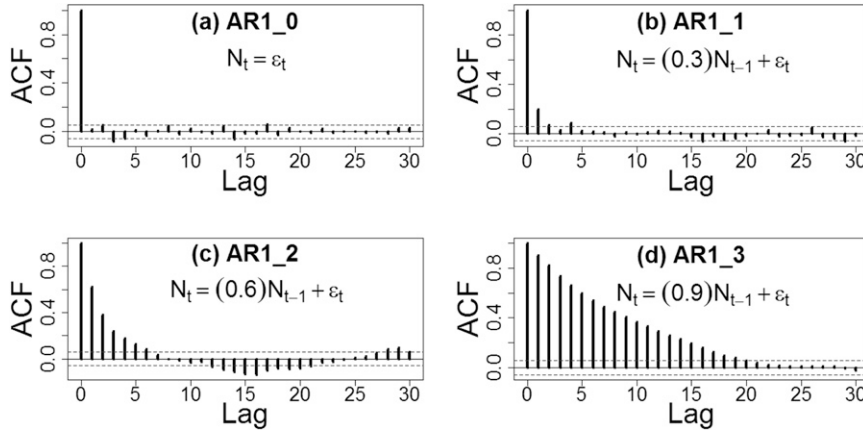


FIG. 2. The ACF of all 100-yr monthly AR1 models with the series number from Table 2.

the number of simulations  $M$  increases. We evaluate the sensitivity of using different numbers of simulations to simulation error (the standard deviation of the average). We would expect simulation error to decrease as the number of simulations increases. Simulation error generally can be calculated from

$$\text{Simulation error} = [\text{Var}(\bar{U})]^{1/2} = \left[ \frac{\text{Var}(U)}{M} \right]^{1/2}. \quad (7)$$

The  $U$  is the value of our simulated variable of interest, which is  $\text{Var}(N)\tau$  in this study. The value  $\bar{U}$  is the average of  $U$  from  $M$  simulations. Figure 4 shows the

sensitivity of the number of simulations to simulation errors (as percent of the  $\bar{U}$ ) for all 100-yr monthly series given in Table 2 using the W98 method. The error converges to the minimum value at a different number of simulations depending on the series. For our study, we can therefore use 10 000 simulations for all series to guarantee convergence to the minimum percent error.

For the L08 method, the criterion for estimating the decorrelation time  $\tau_L = \sum_{\mu=-\infty}^{\infty} \rho_{\mu}$  is to compute the sample autocorrelation coefficients up to  $10 \log_{10}(T)$  lags, then only sum all coefficients of statistically significant lags to get the  $\tau_L$  estimate. Similar to Fig. 4, the sensitivity of the number of simulations to the percent of simulation errors

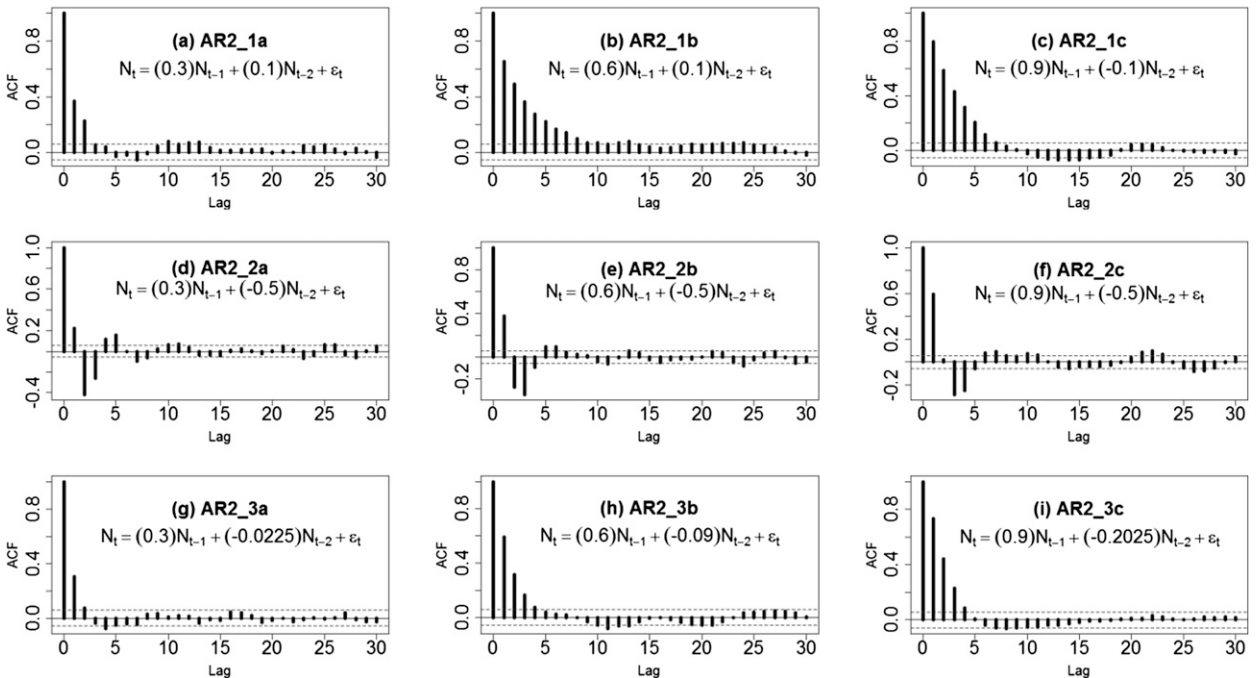


FIG. 3. The ACF of all 100-yr monthly AR2 series with the series number from Table 2.



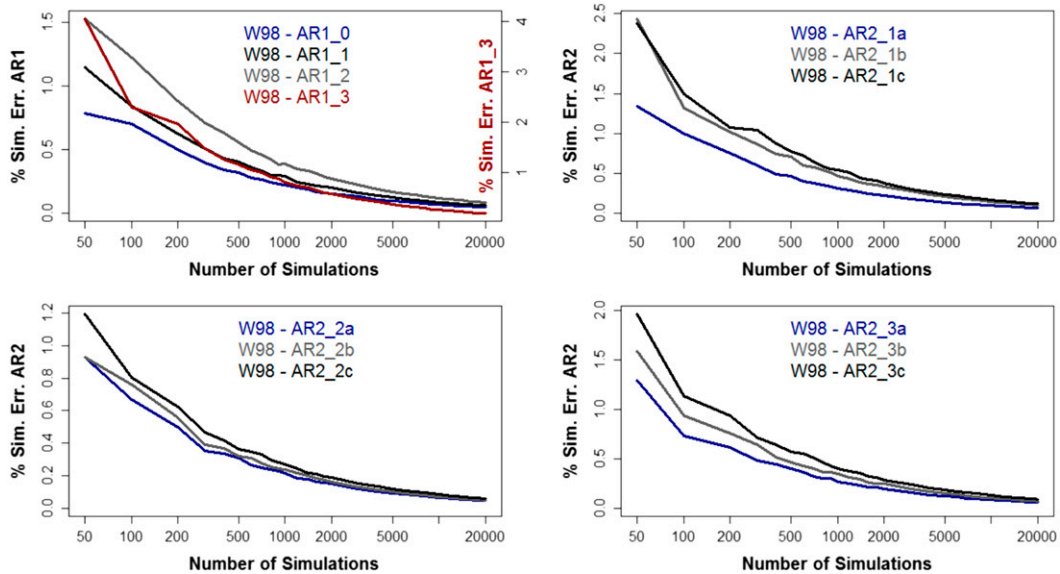


FIG. 4. Sensitivity of the number of simulations to the simulation errors (as percent of its simulated mean) for all 100-yr monthly series (i.e., AR1 and AR2 series) with the W98 method.

for all 100-yr monthly series using the L08 method is shown in Fig. 5. For all series, we therefore use 10 000 simulations to guarantee convergence to the minimum percent error. We notice that similar sensitivity plots (not shown) can be obtained for all 10-yr monthly series using the two methods. Therefore, we set  $M$  (from Fig. 1) to 10 000 simulations for the exploration matrix in Table 2.

*b. Simulation results*

We ran the simulation study according to Fig. 1 using  $M = 10\,000$  simulations for all 13 series. We draw our

conclusion based on comparing each method's  $\text{Var}(N)\tau$  distributions (from 10 000 samples) with respect to the corresponding theoretical values. The simulated performances obtained from the W98 and L08 methods vary depending upon the series. Generally, there are five distinctive performances of the median  $\text{Var}(N)\tau$  with respect to their corresponding theoretical values. Examples of these five distinctive performances are given in Fig. 6. In this figure, box plots of simulated  $\text{Var}(N)\tau$  values at different record lengths as estimated from the two methods are shown. The corresponding theoretical

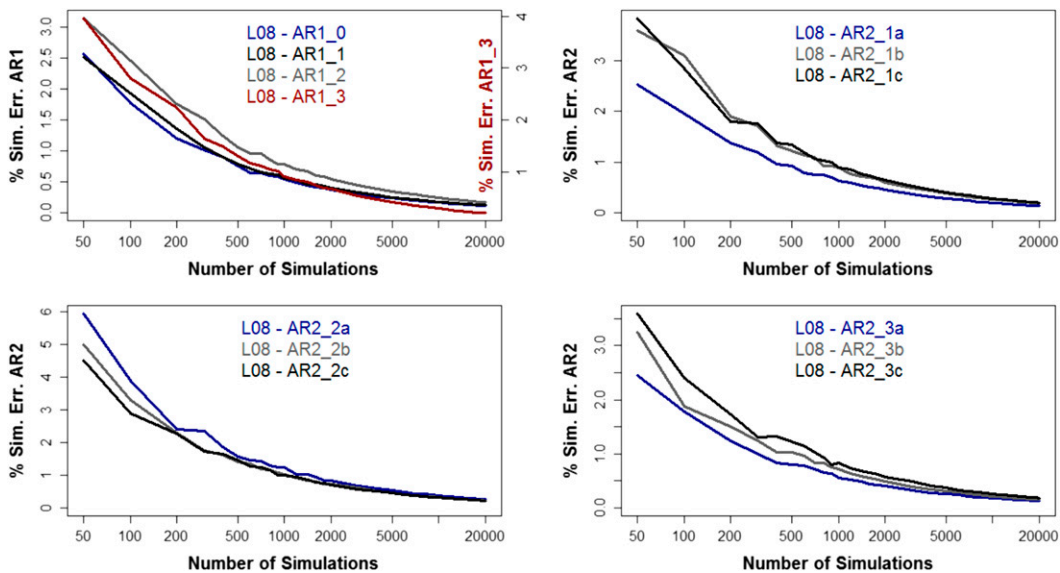


FIG. 5. As in Fig. 4, but for the L08 method.

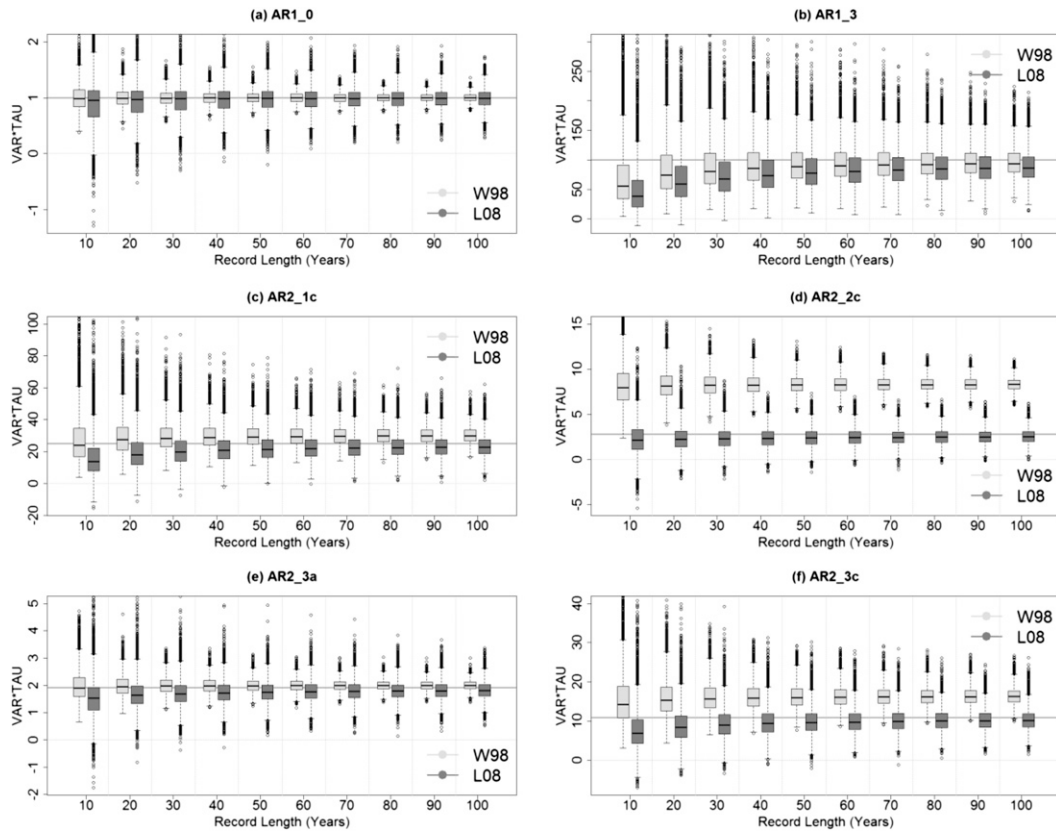


FIG. 6. The box plot of simulated  $\text{Var}(N)\tau$  values lengths as estimated from the two methods for the selective monthly series given in Table 2 at different record for (a) AR1\_0, (b) AR1\_3, (c) AR2\_1c, (d) AR2\_2c, (e) AR2\_3a, and (f) AR2\_3c. The corresponding theoretical  $\text{Var}(N)\tau$  values are given by the dark gray horizontal lines in all figures. The vertical ranges of the figures are not extended out to show the extreme values in order to facilitate readability of more critical information such as the quartile and median values.

$\text{Var}(N)\tau$  values are given by the dark horizontal gray lines. The bottom and top of the box are the first (Q1) and third (Q3) quartiles, and the band inside the box is the second (Q2) quartile (or the median). The interquartile range (IQR) is defined as the difference between Q3 and Q1 (i.e.,  $Q3 - Q1$ ). The low and high ends of the whiskers represent  $Q1 - 1.5\text{IQR}$  and  $Q3 + 1.5\text{IQR}$ , respectively. Any values outside the whiskers represent extreme values. The sample  $\text{Var}(N)$  equation is

$$\frac{\sum_{t=1}^T (N_t - \bar{N})^2}{T - 1},$$

with the numerator being the sum of squares of standard normal deviations, and the denominator is a constant. This sum of squares of deviations follows the chi-square distribution with  $T - 1$  degrees of freedom (Montgomery and Runger 1994). We can conclude that the distribution of simulated  $\text{Var}(N)\tau$  values follows the chi-square

distribution with  $T$  degrees of freedom. The tail regions of the distributions are possible but less probable.

To focus on more critical information such as the quartile and median values, the vertical ranges of the figures are not extended out to the extreme values, which are usually on the tail regions of the distributions. These figures are intended to show the decreasing trend of ranges of  $\text{Var}(N)\tau$  distributions as record lengths increase. Figures 6a and 6b are for AR1\_0 and AR1\_3, respectively. All other AR1 series (AR1\_1 and AR1\_2) perform similarly to these two figures. The AR2\_1c and AR2\_3b series perform similarly and can be represented in Fig. 6c. All the AR2 series with damped sinusoid ACFs (AR2\_2a, AR2\_2b, and AR2\_2c) perform similarly as shown in Fig. 6d. The AR2\_3a perform differently from other series and is shown in Fig. 6e. The AR2\_1a, AR2\_1b, and AR2\_3c series perform similarly as in Fig. 6f.

Table 3 summarizes the simulated performances obtained from the W98 and L08 methods for these five distinctive performances of the median  $\text{Var}(N)\tau$  value.

TABLE 3. Monte Carlo performance comparison for the W98 and L08 methods.

Performance	W98	L08	W98 vs L08	Series no.
Median $\text{Var}(N)\tau$ value	Median closer to its theoretical value as record lengths increase.	Median closer to its theoretical value as record lengths increase.	The W98's medians are closer to the theoretical value than the L08's.	White noise and all AR1 series (Figs. 6a,b)
			The L08's medians are closer to the theoretical value than the W98's, except for the 10-yr record length.	AR2_1a, AR2_1b, and AR2_3c (Fig. 6f)
	Median gradually departing from its theoretical value as record lengths increase.		The L08's medians are closer to the theoretical value than the W98's, except for less than 50-yr lengths.	AR2_1c and AR2_3b (Fig. 6c)
	Median significantly higher than its theoretical value across all lengths.		The W98's medians are closer to the theoretical value than the L08's.	AR2_3a (Fig. 6e)
		The L08's medians are closer to the theoretical value than the W98's.	AR2_2a, AR2_2b, and AR2_2c (Fig. 6d)	
Range of $\text{Var}(N)\tau$ distribution	Smaller ranges as record lengths increase.		The W98's distribution ranges are smaller than the L08's.	White noise, AR1, and AR2 (Fig. 6)
Tail region of $\text{Var}(N)\tau$ distribution	Smaller tail region as record lengths increase.		For the shorter record lengths, only the L08 has tail regions in the negative values because the estimate of $\tau_L$ is occasionally negative.	White noise, AR1, and AR2 (Fig. 6)

The performance of the  $\text{Var}(N)\tau$  distribution ranges and tail regions does not vary with series, but rather vary depending upon the methods. For the AR2 with damped sinusoid ACFs, the L08 gives reasonable medians, while the W98 gives significantly higher medians than the theoretical values. Ranges of the distributions and tail regions from both methods are smaller as record lengths increase. Ranges from the W98 are smaller than those from the L08. The L08's  $\text{Var}(N)\tau$  values occasionally are negative because the sum of all sample autocorrelation coefficients of statistically significant lags for the  $\tau$  estimate is dominated by the negative coefficients. The likelihood (as the percent of total 10 000 simulations) of the L08's negative  $\tau$  value is shown in Fig. 7. As the record length increases, the likelihood becomes smaller. For all series, except the series with damped sinusoid ACFs, the L08 has a likelihood of no more than 3% for the 10-yr record length to give negative  $\tau$  values. For the series with damped sinusoid ACFs (Fig. 7c), the L08 has higher likelihood (as high as 28%) for the 10-yr record length.

Thus far, we have used  $\text{Var}(N)\tau$  as our metric of interest. This metric is not a physical unit that is matter scientifically. As such, we further compute percent

differences for uncertainty of trend estimate  $[\text{Var}(\hat{m})]^{1/2}$  between the simulated medians and the theoretical (truth) for the monthly series as given in Table 2 at selective record lengths of 10, 30, 50, and 100 yr (Fig. 8). General observation for the L08 method is that as record lengths increase, percent differences from the truth decrease for all series. This observation can also be made for all AR1 series, including white noise, estimated by the W98 method (Fig. 8a). For AR2 series estimated by the W98 method (Figs. 8b–d), this observation is true only for the AR2\_1a and AR2\_2b. For all AR1 and white noise series, the W98 method gives smaller percent differences than those from the L08 (Fig. 8a). The maximum percent difference from the theoretical (truth) across all record lengths for the white noise monthly series is 1% from the W98 and 2% from the L08. The AR1 series has a maximum difference of 26% using the W98 and 38% using the L08. Therefore, the W98 is preferred over the L08 for all AR1 and white noise series. For AR2 series (Figs. 8b,d), the L08 gives smaller percent differences than those from the W98 for record lengths of more than 10 yr but gives larger differences for a record length of 10 yr. Maximum differences are



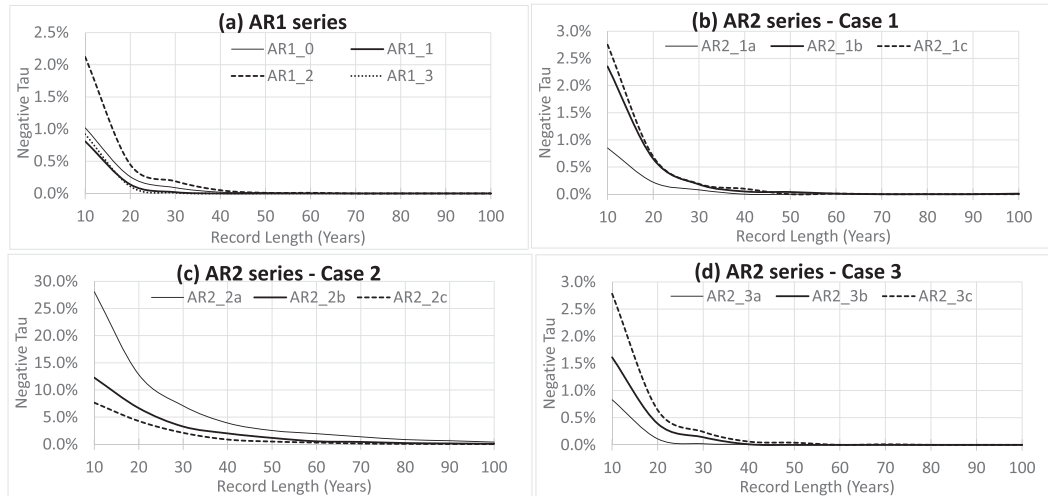


FIG. 7. The likelihood of negative  $\tau$  values from the L08 for all simulated series, based on the percent total of the 10000 computer simulations: (a) AR1 series; (b) AR2 series with ACF exhibiting a mixture of two exponential decay patterns (case 1, AR2\_1); (c) AR2 series with ACF exhibiting damped sinusoid pattern (case 2, AR2\_2); and (d) AR2 series with ACF exhibiting a single exponential decay pattern (case 3, AR2\_3).

17% from the W98 and 26% from the L08 for AR2 series with a mixture of two exponential decay ACF patterns (case 1 in Fig. 8b). Maximum differences are 22% from the W98 and 20% from the L08 for the AR2 series with single exponential decay ACF pattern (case 3 in Fig. 8d). For the AR2 series with damped sinusoid ACFs (AR2 case 2 in Fig. 8c), the W98 gives a maximum of 73% differences for the AR2 series, while the L08 gives a maximum of 30%.

### 3. Trend analysis for Clouds and the Earth's Radiant Energy System

CERES instruments measure broadband radiances in three channels: the shortwave (wavelength from 0.2 to  $5 \mu\text{m}$ ), the longwave (wavelength from 5 to  $50 \mu\text{m}$ ), and the total (wavelength from 0.2 to  $50 \mu\text{m}$ ). Observed radiances are converted to TOA irradiances using angular distribution models (ADMs) (Loeb et al. 2001, 2005). A

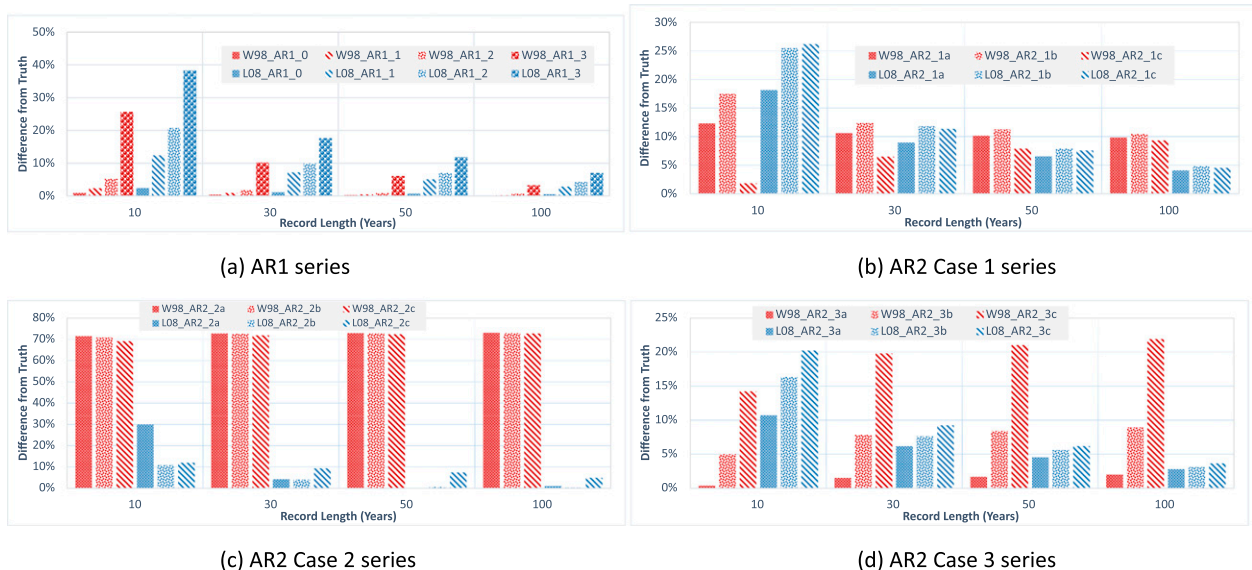


FIG. 8. Percent differences for uncertainty of trend estimate  $[\text{Var}(\hat{m})]^{1/2}$  between the simulated median and the theoretical (truth) for the monthly series as given in Table 2 at record lengths of 10, 30, 50, and 100 yr: (a) AR1 series; (b) AR2 series with ACF exhibiting a mixture of two exponential decay patterns (case 1, AR2\_1); (c) AR2 series with ACF exhibiting damped sinusoid pattern (case 2, AR2\_2); and (d) AR2 series with ACF exhibiting a single exponential decay pattern (case 3, AR2\_3).

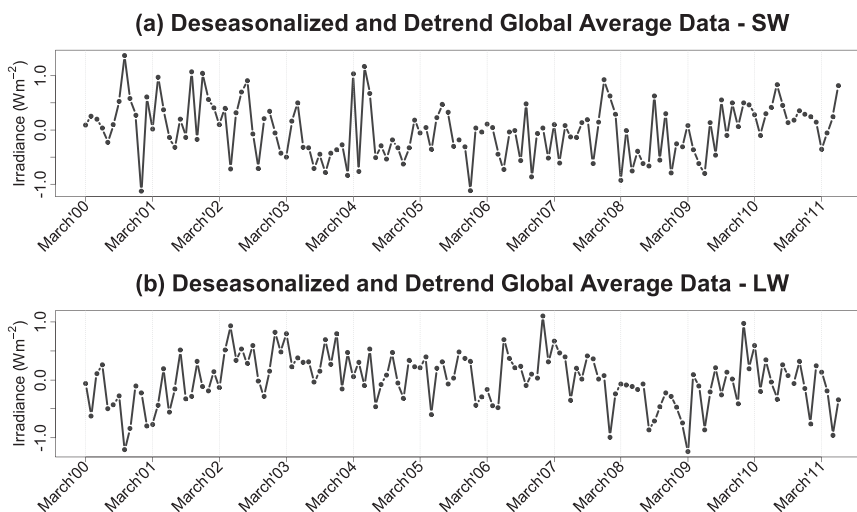


FIG. 9. Time series plots of monthly deseasonalized and detrended global averages for (a) reflected shortwave irradiances and (b) emitted longwave irradiances.

detailed description of CERES and its instruments is provided in Wielicki et al. (1996).

The 136 months of reflected shortwave and emitted longwave irradiances at TOA estimated from observed CERES broadband radiances from March 2000 through June 2011 [Energy Balanced and Filled (EBAF) edition 2.6r; Loeb et al. 2009] are used in this study. The reflected shortwave and emitted longwave irradiances monthly series are globally averaged and deseasonalized. No statistically significant trend was found in either deseasonalized series. This is an indication that the current record length may be insufficient to detect any irradiance change. Figure 9 shows time series plots for deseasonalized and detrended global average reflected shortwave and longwave data. The standard deviations  $\text{Std}(N_t)$  for these series are 0.51 and  $0.45 \text{ W m}^{-2}$  for shortwave and longwave, respectively. These values are consistent with those reported in Kato (2009).

Figure 10 gives autocorrelation functions (Figs. 10a,b) for both monthly series. Figures 10c and 10d provide autocorrelation functions for both monthly series if they were to behave as the perfect AR1 series, using the estimated lag-1 sample autocorrelation coefficient from the shortwave (Fig. 10c) and longwave (Fig. 10d). The ACFs of both series (Figs. 10a,b) do not exhibit purely damped sinusoid.

Using Eqs. (3) and (2), the uncertainties of trend estimate for the shortwave series are 1.27 and  $1.48 \text{ mW m}^{-2}$  for the W98 and L08, respectively. For the longwave series, the uncertainties of trend estimate for the longwave are 1.40 and  $2.31 \text{ mW m}^{-2}$  for the W98 and L08, respectively. Uncertainty of trend estimate for the longwave series is higher than that for the shortwave series. By

examining Figs. 10a and 10b, the longwave irradiance series has more significant lags (at a 95% confidence level) than the shortwave series. Consequently, the longwave series takes a longer time to have monthly data uncorrelated (i.e., larger decorrelation time), and, in turn, its estimated uncertainty is higher than that of the shortwave series.

The shortwave ACF looks similar to the perfect AR1 (Figs. 10a,c). From section 2, the simulations of the AR1 series for 10-yr record length (Fig. 8a) have shown that the W98 method produces better uncertainty (in terms of smaller percent differences from the theoretical value) than that from the L08 method. Therefore, we recommend using the W98 method for the shortwave series. The longwave ACF does not exhibit a purely exponential decay nor damped sinusoid pattern, but rather a mixture of both patterns, making the exponential decay slower than a purely exponential decay pattern. This behavior can be observed by comparing Figs. 10b and 10d. The W98, which is based on exponential decay of the ACF pattern similar to Fig. 10d, underestimates the uncertainty of trend estimate for the longwave series with its ACF as in Fig. 10b, while the L08 provides a more reasonable value. Therefore, we recommend using the L08 method for the longwave series.

Another metric of interest is the number of years  $n$  needed to detect a trend (per year) of magnitude  $|\omega_o|$  with 90% testing power,<sup>1</sup> as estimated by the W98:

<sup>1</sup> The testing power is the power to detect correctly 90% of the time a trend that is at least twice as large as its uncertainty (i.e.,  $|\omega_o|/\text{Std}(\omega_o) > 2$ ).

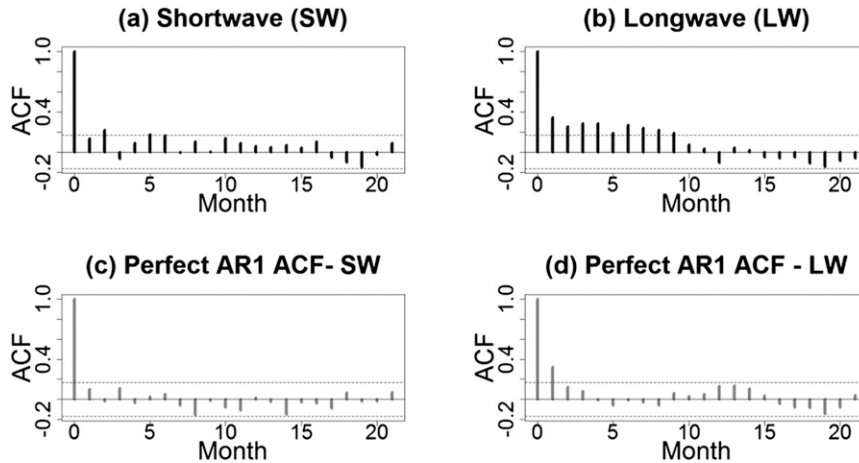


FIG. 10. Autocorrelation functions of deseasonalized and detrended monthly series for (a) reflected shortwave irradiances and (b) emitted longwave irradiances. Autocorrelation functions if the (c) reflected shortwave and (d) emitted longwave irradiances were a perfect AR1 time series using lag-1 sample autocorrelation coefficients estimated from the corresponding series.

$$n = \left[ \frac{3.3 \sqrt{\text{Var}(N)\tau_W}}{|\omega_o|} \right]^{2/3}. \tag{8}$$

The number of years estimated by the L08 is similar to Eq. (8), with  $\text{Var}(N)\tau_L$  replacing  $\text{Var}(N)\tau_W$ . Figure 11 graphs the estimates of time (in years) to detect the various  $|\omega_o|/\text{Std}(N_t)$  ratios. As the trend magnitude (or the ratio) gets smaller, the time to detect the smaller magnitude takes longer. For the reflected shortwave (Fig. 11a), there is a slight difference between the W98 and L08 to detect various ratios, with a maximum difference of about 6 yr at the ratio of 0.01. For the longwave (Fig. 11b), there is a noticeable difference between the two methods to detect various ratios, with a maximum difference of about 14 yr at the ratio of 0.01.

As an example, based on our earlier recommendation of using the W98 to estimate the uncertainty for the reflected shortwave irradiances, if the trend per year is about  $5.06 \text{ mW m}^{-2}$  [i.e.,  $\text{Std}(N_t)$  of  $0.51 \text{ W m}^{-2}$  and the  $|\omega_o|/\text{Std}(N_t)$  ratio of 0.01], it will take about 52.3 yr to detect it using the W98 method. For the emitted longwave irradiances, based on our earlier recommendation of using the L08, if the trend per year is about  $4.49 \text{ mW m}^{-2}$  [i.e.,  $\text{Std}(N_t)$  of  $0.45 \text{ W m}^{-2}$  and the ratio of 0.01], it will take about 84.4 yr to detect it using the L08 method.

#### 4. Conclusions

In this paper, two trend analysis methods, the methods of W98 and L08, are compared. The two methods are

based on linear regression models with covariance structures of the time series data properly accounted for under different underlying assumptions. The W98 assumes that the time series data follow the first-order autoregressive model (AR1), where its autocorrelation function (ACF) decays exponentially with lags. In practice, only the lag-1 sample autocorrelation coefficient is used to estimate the uncertainty of the trend estimate making it simple to use. However, with any departure from the exponential decay pattern, the W98 is expected to perform poorly.

The L08 method, on the other hand, does not assume any autoregressive model or any correlation structure. It uses theoretical autocorrelation coefficients at all lags (i.e.,  $\sum_{\mu=-\infty}^{\infty} \rho_{\mu}$ , where  $\rho_{\mu}$  is the lag  $\mu$  autocorrelation coefficient) in the analysis making it ideally versatile. In practice, the theoretical coefficients  $\rho_{\mu}$  of the time series are not known, but can be estimated by its sample autocorrelation coefficients  $\hat{\rho}_{\mu}$  from the available time series data of length  $T$  months. Taking sample coefficients at all available lags in estimating L08's uncertainty of the trend will provide zero uncertainty; specifically,  $\sum_{\mu=-(T-1)}^{T-1} \hat{\rho}_{\mu} = 0$  is zero. Therefore, it is necessary to set up the criterion to determine which sample autocorrelation coefficients to include in the computation. The criterion used in this study is to include only the statistically significant coefficients up to  $10 \log_{10}(T)$  lags in the summation.

Because the  $\text{Var}(N)\tau$  term is directly related to the uncertainty of trend estimate [Eqs. (2) and (3)] from the two methods, comparing this term estimated by the two methods is adequate in a Monte Carlo simulation study. The comparison is performed by the simulation study with the theoretically and randomly generated theoretical

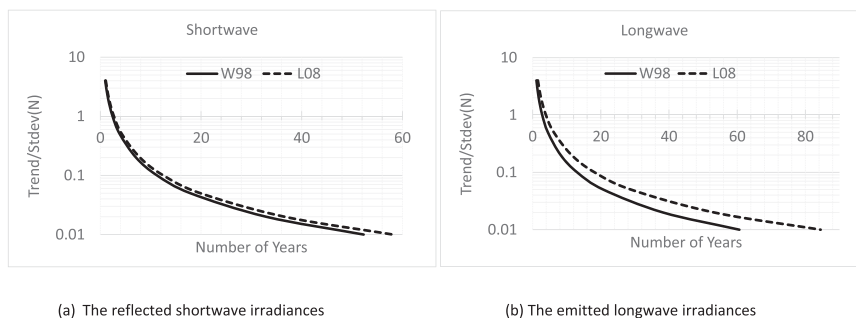


FIG. 11. Number of years required to detect various  $|\omega_o|/\text{Std}(N_t)$  ratios with 90% testing power (the power to detect correctly 90% of the time a trend that is at least twice as large as its uncertainty) by the W98 and L08 methods: (a) the reflected shortwave irradiances and (b) the emitted longwave irradiances.

AR1 time series with  $\rho_1$  of 0.0, 0.3, 0.6, and 0.9, representing white noise, low-, medium-, and high-correlated series. The study also extends to the theoretically and randomly generated second-order autoregressive (AR2) series. We limit our simulation study to AR2 for simplicity because it already captures phenomena that are likely to complicate the W98.

For the AR2 with damped sinusoid ACFs, the L08 gives reasonable medians, while the W98 gives significantly higher medians than the theoretical values. The ranges of  $\text{Var}(N)\tau$  distributions and tail regions of the distributions from both methods are smaller as record lengths increase. Ranges are smaller when the W98 is used. The L08 occasionally gives a negative  $\tau$  value, which, in turn, makes a negative  $\text{Var}(N)\tau$  value. For the shorter record lengths, the L08's tail regions are negative in value; however, the negative tail regions become smaller as the record lengths increase.

From the Monte Carlo simulation study, the white noise monthly series has a maximum percent difference from the theoretical (truth) of 1% and 2% from the W98 and L08, respectively. The AR1 series has a maximum difference of 26% using the W98 and 38% using the L08. The AR2 series with a mixture of two exponential decay ACF patterns has a maximum difference of 17% using the W98 and of 26% using the L08. The AR2 series with single exponential decay ACF pattern has a maximum difference of 22% using the W98 and of 20% using the L08. The AR2 series with the damped sinusoid ACF pattern has a maximum of 73% using the W98 and 30% using the L08.

Both methods are then applied to the time series of deseasonalized anomalies derived from monthly and globally averaged reflected shortwave and emitted longwave irradiances at the top of the atmosphere (TOA), as estimated from CERES broadband radiances measurements during March 2000 through June 2011. Examination of the autocorrelation structure shows that the reflected

shortwave series has an exponential decay pattern, while the longwave series has a mixture of exponential decay and damped sinusoid patterns. Therefore, we recommend using the W98 with the reflected shortwave series and the L08 with the longwave series for trend analysis. The uncertainty of trend estimate for the shortwave series is approximately  $1.27 \text{ mW m}^{-2}$ . Its month-to-month standard deviation is  $0.51 \text{ W m}^{-2}$ , and if the trend per year is 100 times smaller than this standard deviation (i.e., trend per year of approximately  $0.0051 \text{ W m}^{-2}$ ), it will take about 52.3 yr to detect it. For the longwave series, the uncertainty of the trend estimate is about  $2.31 \text{ mW m}^{-2}$ . Its month-to-month standard deviation is  $0.45 \text{ W m}^{-2}$ , and if the trend per year is 100 times smaller than this standard deviation (i.e., trend per year of approximately  $0.0045 \text{ W m}^{-2}$ ), it will take about 84.4 yr to detect it. Note that while the results in this paper are applied to radiation data as an example, they apply in general to all climate change time series.

One potential for future study is to explore other more accurate methods for estimating  $\sum_{\mu=-\infty}^{\infty} \hat{\rho}_{\mu}$  for the L08. One suggested method is to fit an autoregressive moving average (ARMA) model and use its model parameter estimates for  $\hat{\rho}_{\mu}$  estimation. Some bias correction methods (Lee and Lund 2004) could be used to substantially increase the accuracy of the  $\hat{\rho}_{\mu}$  estimation.

We have shown that estimates of climate trend uncertainties can be sensitive to the methods used. The tests performed in this paper can be applied to any climate time series to determine if the results are sensitive to the method used, either the W98 or L08. If the results are sensitive, then the choice of the method should be based on the physical nature shown in the autocorrelation structure of the climate time series.

*Acknowledgments.* We thank Dr. Stephen Leroy for useful discussions and all reviewers' comments for valuable insights. We also thank Ms. Amber Richards and

Dr. Joe A. Walker for proof reading the manuscript. This work is supported by the NASA CLARREO project.

of the trend per year,  $N_t$  is the unexplained noise of the measurement data, and  $t = 1, 2, 3, \dots, T$  months. It is assumed that the noise  $N_t$  follows an AR2 model with

APPENDIX

**Derivation for the AR2 Theoretical Value Given in Section 2a**

The derivation of uncertainty of trend estimate for the AR2 model is an extension of the AR1's derivation as in Tiao et al. (1990). Let  $Y_t$  be the measurement of interest. The linear trend model assumes

$$Y_t = \alpha + S_t + \omega X_t + N_t, \tag{A1}$$

where  $\alpha$  is a constant term,  $S_t$  is the seasonal component,  $X_t = t/12$  is the linear trend function,  $\omega$  is the magnitude

$$N_t = \phi_1 N_{t-1} + \phi_2 N_{t-2} + \varepsilon_t, \tag{A2}$$

where  $\phi_1$  and  $\phi_2$  are parameters of the model. From Eq. (A1), we can obtain

$$\phi_1 Y_{t-1} = \phi_1 \alpha + \phi_1 S_{t-1} + \omega \phi_1 X_{t-1} + \phi_1 N_{t-1}, \tag{A3}$$

and

$$\phi_2 Y_{t-2} = \phi_2 \alpha + \phi_2 S_{t-2} + \omega \phi_2 X_{t-2} + \phi_2 N_{t-2}. \tag{A4}$$

From Eqs. (A1), (A3), and (A4),

$$(Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2}) = \alpha(1 - \phi_1 - \phi_2) + (S_t - \phi_1 S_{t-1} - \phi_2 S_{t-2}) + \omega(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}) + (N_t - \phi_1 N_{t-1} - \phi_2 N_{t-2}). \tag{A5}$$

Replacing the last term on the right-hand side with  $\varepsilon_t$  [from Eq. (A2)] and  $X_t = t/12$  in Eq. (A5), we get

$$(Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2}) = \alpha(1 - \phi_1 - \phi_2) + (S_t - \phi_1 S_{t-1} - \phi_2 S_{t-2}) + \omega \frac{[t - \phi_1(t-1) - \phi_2(t-2)]}{12} + \varepsilon_t. \tag{A6}$$

Letting  $Y_t^* = (Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2})$ ,  $\alpha^* = \alpha(1 - \phi_1 - \phi_2)$ ,  $S_t^* = (S_t - \phi_1 S_{t-1} - \phi_2 S_{t-2})$ , and  $t^* = [t - \phi_1(t-1) - \phi_2(t-2)]/12$ , Eq. (A6) becomes

$$Y_t^* = \alpha^* + S_t^* + \omega t^* + \varepsilon_t. \tag{A7}$$

The estimate of the regression coefficient  $\omega$  can be obtained essentially by ordinary least squares calculations applied to Eq. (A7). Therefore,

$$\hat{\omega} = \frac{\sum (t^* - \bar{t}^*) Y_t^*}{\sum (t^* - \bar{t}^*)^2}.$$

Because  $t^* = [t - \phi_1(t-1) - \phi_2(t-2)]/12 = [t(1 - \phi_1 - \phi_2) + (\phi_1 + \phi_2)]/12$  and  $\bar{t}^* = [\bar{t}(1 - \phi_1 - \phi_2) + (\phi_1 + \phi_2)]/12$ , therefore  $t^* - \bar{t}^* = [(t - \bar{t})(1 - \phi_1 - \phi_2)]/12$ . As such,

$$\begin{aligned} \hat{\omega} &= \frac{\sum (t - \bar{t}) Y_t^* (1 - \phi_1 - \phi_2)/12}{\sum (t - \bar{t})^2 (1 - \phi_1 - \phi_2)^2 / 12^2} \\ &= \frac{12}{(1 - \phi_1 - \phi_2)} \frac{\sum (t - \bar{t}) Y_t^*}{\sum (t - \bar{t})^2}. \end{aligned} \tag{A8}$$

And from simple linear regression of Eq. (A7),

$$\text{Var}(\hat{\omega}) = \frac{\text{Var}(\varepsilon_t)}{\sum (t^* - \bar{t}^*)^2} = \frac{\text{Var}(\varepsilon_t)}{\sum (t - \bar{t})^2 (1 - \phi_1 - \phi_2)^2 / 12^2}. \tag{A9}$$

Because  $\sum_{t=1}^T (t - \bar{t})^2 = T(T^2 - 1)/12$ , Eq. (A9) becomes

$$\begin{aligned} \text{Var}(\hat{\omega}) &= \frac{12^3 \text{Var}(\varepsilon_t)}{T(T^2 - 1)(1 - \phi_1 - \phi_2)^2} \\ &\approx \frac{\text{Var}(\varepsilon_t)}{(T/12)^3 (1 - \phi_1 - \phi_2)^2}. \end{aligned} \tag{A10}$$

Because  $\omega$  is trend per year, the variance of trend per month  $m$  can be computed from

$$\text{Var}(\hat{m}) = \text{Var}\left(\frac{\hat{\omega}}{12}\right) = \frac{\text{Var}(\hat{\omega})}{12^2} \approx \frac{12}{T^3} \frac{\text{Var}(\varepsilon_t)}{(1 - \phi_1 - \phi_2)^2}. \tag{A11}$$

Comparing Eq. (A11) with Eqs. (2) or (3), we get



$$\text{Var}(N_i)\tau = \frac{\text{Var}(\varepsilon_i)}{(1 - \phi_1 - \phi_2)^2}. \quad (\text{A12})$$

## REFERENCES

- Hassani, H., 2010: A note on the sum of the sample autocorrelation function. *Physica A*, **389**, 1601–1606, doi:10.1016/j.physa.2009.12.050.
- Hill, W. J., P. N. Sheldon, and J. J. Tiede, 1977: Analyzing worldwide total ozone for trends. *Geophys. Res. Lett.*, **4**, 21–24, doi:10.1029/GL004i001p00021.
- Kato, S., 2009: Interannual variability of the global radiation budget. *J. Climate*, **22**, 4893–4907, doi:10.1175/2009JCLI2795.1.
- Lee, J., and R. Lund, 2004: Revisiting simple linear regression with autocorrelated errors. *Biometrika*, **91**, 240–245, doi:10.1093/biomet/91.1.240.
- Leroy, S. S., J. G. Anderson, J. Dykema, and R. Goody, 2008a: Testing climate models using thermal infrared spectra. *J. Climate*, **21**, 1863–1875, doi:10.1175/2007JCLI2061.1.
- , —, and G. Ohring, 2008b: Climate signal detection times and constraints on climate benchmark accuracy requirements. *J. Climate*, **21**, 841–846, doi:10.1175/2007JCLI1946.1.
- Loeb, N. G., K. J. Priestley, D. P. Kratz, E. B. Geier, R. N. Green, B. A. Wielicki, P. O. R. Hinton, and S. K. Nolan, 2001: Determination of unfiltered radiances from the Clouds and the Earth's Radiant Energy System (CERES) instrument. *J. Appl. Meteor.*, **40**, 822–835, doi:10.1175/1520-0450(2001)040<0822:DOURFT>2.0.CO;2.
- , S. Kato, K. Loukachine, and N. M. Smith, 2005: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the *Terra* satellite. Part I: Methodology. *J. Atmos. Oceanic Technol.*, **22**, 338–351, doi:10.1175/JTECH1712.1.
- , B. A. Wielicki, D. R. Doelling, G. L. Smith, D. K. Keyes, S. Kato, N. Manalo-Smith, and T. Wong, 2009: Toward optimal closure of the earth's top-of-atmosphere radiation budget. *J. Climate*, **22**, 748–766, doi:10.1175/2008JCLI2637.1.
- Montgomery, D. C., and G. C. Runger, 1994: *Applied Statistics and Probability for Engineers*. John Wiley and Sons, 895 pp.
- , C. L. Jennings, and M. Kulahci, 2008: *Introduction to Time Series Analysis and Forecasting*. John Wiley and Sons, 441 pp.
- Reinsel, G. C., G. C. Tiao, J. J. DeLuisi, S. Basu, and K. Carriere, 1989: Trend analysis of aerosol-corrected Umkehr ozone profile data through 1987. *J. Geophys. Res.*, **94**, 16 373–16 386, doi:10.1029/JD094iD13p16373.
- Ross, S. M., 2006: *Simulation*. 4th ed. Elsevier, 298 pp.
- Tiao, G. C., and Coauthors, 1990: Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation. *J. Geophys. Res.*, **95**, 20 507–20 517, doi:10.1029/JD095iD12p20507.
- Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S*. 4th ed. Springer-Verlag, 495 pp.
- Von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Weatherhead, E. C., and Coauthors, 1998: Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J. Geophys. Res.*, **103**, 17 149–17 161, doi:10.1029/98JD00995.
- Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System experiment. *Bull. Amer. Meteor. Soc.*, **77**, 853–868, doi:10.1175/1520-0477(1996)077<<0853:CATERE>>2.0.CO;2.