Towards a robust test on North America warming trend and precipitable water content increase

Jih-Wang Wang,[1] Ke Wang,[2] Roger A. Pielke Sr.,[3] John C. Lin,[4] and Toshihisa Matsui[5,6]

[1]   An increase in the atmospheric moist content has been generally assumed when the lower-tropospheric temperature (T_{col}) increases, with relative humidity holding steady. Rather than using simple linear regression, we propose a more rigorous trend detection method that considers time series memory. The autoregressive moving-average (ARMA) parameters for the time series of T_{col}, precipitable water vapor (PWAV), and total precipitable water content (PWAT) from the North American Regional Reanalysis data were first computed. We then applied the Monte Carlo method to replicate the ARMA time series samples to estimate the variances of their Ordinary Least Square trends. Student's t tests showed that T_{col} from 1979 to 2006 increased significantly; however, PWAV and PWAT did not. This suggests that atmospheric temperature and water vapor trends do not follow the conjecture of constant relative humidity over North America. We thus urge further evaluations of T_{col}, PWAV, and PWAT trends for the globe. **Citation:** Wang, J.-W., K. Wang, R. A. Pielke Sr., J. C. Lin, and T. Matsui (2008), Towards a robust test on North America warming trend and precipitable water content increase, *Geophys. Res. Lett.*, *35*, L18804, doi:10.1029/2008GL034564.

1.   Introduction

[2]   Atmospheric reanalysis data have been used for temperature trend detection [e.g., *Pielke et al.*, 1998; *Chelliah and Ropelewski*, 2000; *Chase et al.*, 2000]. When considering water vapor in the atmosphere, the Clausius-Clapeyron (C-C) equation has been proposed to explain a significant portion of the temporal and spatial covariation between near-surface air temperature and column-integrated precipitable water vapor content, and a near constant global average relative humidity has been obtained or presumed in global model simulations of the relationship between temperature increase and relative humidity [e.g., *Held and Soden*, 2006; *Soden and Held*, 2006]. For an open system (e.g., regional domain), however, the implication of the C-C equation and fixed relative humidity needs further examination.

[3]   The evaluation of water vapor trend is critical for the Earth's radiation budget due to its strong greenhouse radiative forcing, which is estimated to be 14 times larger than CO_2 [*Zastawny*, 2006]. As clouds (water condensate) may either warm or cool the Earth surface by differentially affecting shortwave and longwave radiation, clouds become the greatest source of uncertainty in this estimate, since the frequency and types of clouds are not well observed or simulated [*Arakawa*, 2004]. Temporal variations in water vapor and condensate are, therefore, worthy of more investigation.

[4]   One problem in detecting the trend in these data is their variations in time and space. *Box and Jenkins* [1976] presented a widely used methodology to account for the autocorrelation of time series, called the autoregressive moving average (ARMA) model. However, trend detection still remains a difficult problem because the distribution of trend estimates is not random (white noise) when memory exists in the time series.

[5]   In this study, a new methodology is proposed for robust trend detection based on the work of *Box and Jenkins* [1976] (Section 2). In previous studies, *Weatherhead et al.* [1998] developed a revised formula for trend variability considering an AR(1) process (autoregressive process of order one; see next section for more details) for the time series. Their exact form of standard deviation of the trend, however, requires another substantial revision if more ARMA parameters are to be applied; see *Fomby and Vogelsang* [2002] and *Woodward and Gray* [1993] for more examples.

[6]   Due to the complexity of an algebraic expression of trend variance, we conducted Monte Carlo experiments to empirically obtain values for trend variance (Section 3). Our approach obtains a trend variance from a simulated population instead of estimating it based on one single time series sample.

[7]   Our paper focuses on analyzing the domain-averaged temporal variability of lower-tropospheric temperature (T_{col}), column-integrated precipitable water vapor content (PWAV), and column-integrated total precipitable water content (PWAT, including water vapor *and* condensates) by using the North American Regional Reanalysis (NARR) data [*Mesinger et al.*, 2006]. We then test the hypothesis whether an increasing (decreasing) trend of atmospheric water vapor and a warming (cooling) trend of atmospheric temperature over North America exist, as in the general expectation in the science community [e.g., *Core and Extended Writing Team*, 2007].

[1]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado, USA.
[2]Department of Statistics, Colorado State University, Fort Collins, Colorado, USA.
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.
[4]Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, Ontario, Canada.
[5]Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA.
[6]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

2. Statistical Model

2.1. Autoregressive Moving Average (ARMA) Analysis

[8] For estimating the temporal trend of a geophysical parameter, one can use the linear regression model:

$$y_t = \beta_1 + \beta_2 t + u_t, \tag{1}$$

where y_t is the observation value at time t, β_1 is the intercept of the regressed line, β_2 is the slope, and u_t is the residual. β_1 may contain the information of known regular period, such as daily or seasonal cycles. If the time series has memory, it is widely known that the t statistic for the estimated trend ($\hat{\beta}_2$) *cannot* be constructed by treating u_t as independent and identically-distributed (i.i.d.) random variables, and instead u_t is better modeled as an ARMA(p, q) process [*Box and Jenkins*, 1976]:

$$u_t = \sum_{i=1}^{p} a_i u_{t-1} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t, \tag{2}$$

where $a_1, \ldots, a_p$ are the autoregressive parameters, $b_1, \ldots, b_q$ are moving-average parameters, and ε_t is a white noise process with zero mean and variance σ^2. The subscript t, $t - i$, and $t - j$ denote a specific time, while p and q represent the number of previous residual and white noise that significantly influence the present residual. If $\{y_t\}$ is a Gaussian time series, the Generalized Least Square (GLS) or Ordinary Least Square (OLS) estimator of β_2 is also normally distributed since they are a linear function of y_t.

2.2. Estimating the Variance of $\hat{\beta}_2$

[9] For the correlated random variables u_t, the GLS estimator for β_2 can be obtained, and the test for trend easily follows if the covariance matrix were known. *Weatherhead et al.* [1998] derived the exact form of the covariance matrix for an AR(1) process. Unfortunately, the analytical form of the error-covariance matrix is hard to obtain for a more general ARMA(p, q) process. Therefore, we use the Monte Carlo method to generate a large number of realizations of u_t, assuming a null trend in the first place, and compute $Var(\hat{\beta}_2)$ directly. The purpose is to perform a Student's t test on the North America T_{col} and PWAV/PWAT trends. See Section 3.2 for further details.

3. Trend Detection

3.1. Data Description

[10] The North American Regional Reanalysis (NARR), provided by the National Centers for Environmental Prediction (NCEP; http://www.emc.ncep.noaa.gov/mmb/rreanl/) is used for our assessment of the trends. NARR has considerable improvements in both resolution and accuracy from the global reanalysis, because of its assimilation of observed precipitation and 10 meter winds, and its enhanced temporal and spatial resolution [*Mesinger et al.*, 2006]. It provides information that is as accurate as any available to evaluate trends in lower-tropospheric temperature and water vapor in this region.

[11] The grid spacing in the NARR is about 32 km in a Lambert-Conformal projection. The TIROS Operational Vertical Sounder (TOVS) Level 1B radiance data is assimilated into the reanalysis model to cover the oceans for precipitable water vapor content. Note that in NARR, water vapor and condensates may be added into the atmospheric column if PWAT is less than the observed surface precipitation.

[12] We utilize the monthly PWAV, PWAT, surface pressure, and 500 mb geopotential height from May 1979 to January 2006 in this research. We found that the PWAT in April 1979 and most of the PWAV throughout 2006 were apparently outliers, and thus we removed these time periods.

[13] NARR does not supply PWAV directly; thus we integrate specific humidity from the surface to the top of the reanalysis domain. We trim the original NARR domain along the north boundary by four grid points, the east boundary by 14 grid points, and the west boundary by 60 grid points due to some missing values. This removes 23% of the NARR data from the analysis, with most of the removal over the Pacific Ocean. The domain that we investigated thus covers the whole of North America, part of the Atlantic Ocean, Pacific Ocean, and Greenland.

[14] We define lower-tropospheric temperature (T_{col}, surface-to-500mb column mean temperature), following the hydrostatic approximation and ideal gas law, as:

$$dp = -\rho g dz = -\frac{p}{RT} g dz$$

$$\Rightarrow T_{col} = \overline{T}_{sfc-500mb} = \frac{g}{R}(H_{500mb} - z_{sfc})/\ln(p_{sfc}/500mb)$$

Monthly domain-averaged T_{col}, PWAV, and PWAT were calculated so as to acquire three time series with 321 elements in each. Figure 1 shows the temporal variation of these three time series after removing the seasonal cycle, together with their OLS linear fit. The seasonal cycle was composed of 12 monthly values, which were computed as 26- or 27-years averages. The trend for T_{col} is +0.248 degree per decade, for PWAV +0.00619 kg/m^2 per decade, and for PWAT +0.0108 kg/m^2 per decade.

3.2. Experiment Procedure

[15] After removing the seasonality of the time series, we a) used the OLS method to estimate the trend of the time series, b) subtracted the OLS trend from the time series, c) estimated ARMA(p, q) parameters, and d) generated 100,000 time series realizations using the ARMA(p, q) parameters. To obey the hypothesis of null trend, the time series realizations were generated following equations (1) and (2), given $\beta_2 = 0$.

3.2.1. ARMA Parameters Estimation

[16] T_{col}, PWAV, and PWAT time series from NARR were first assessed to be Gaussian by checking "normal QQ" plots. The normal QQ plot is a graphical display of how well the normal distribution describes the data [*Chambers et al.*, 1983]. The values of p and q were then determined by the minimum AIC (Akaike Information Criteria), which is a measure of the goodness of fit of an estimated statistical model [*Akaike*, 1973], and a preliminary estimation was conducted by the Hannan-Rissanen procedure [*Hannan and Rissanen*, 1982]. Once we specify values of p and q and the initial values for the parameter estimates, the maximum likelihood estimates were obtained as the final estimates for a_i and b_j. The whole estimation procedure was implemented with the software package ITSM2000 (a new windows release of the ITSM software on CD that accompanies the fourth printing of *Brockwell and Davis* [2000]). The com-
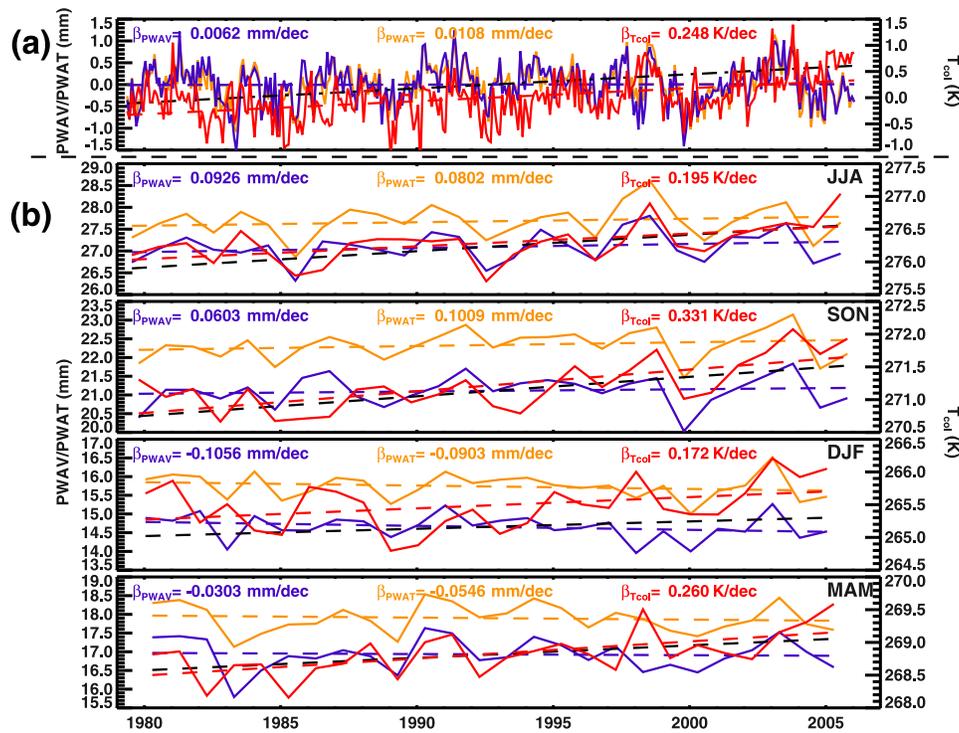
Figure 1. (a) The North American domain-averaged de-seasoned PWAV (blue), PWAT (brown), and T_{col} (red), monthly anomaly time series. (b) The North American domain-averaged PWAV, PWAT, and T_{col} yearly time series by season. The dashed lines represent OLS linear fit, and the magnitude of the trends are also shown. The black dashed lines indicate PWAV trends for a fixed relative humidity scenario.

puted values are listed in Table 1. Note that (p, q) is (2, 2) for T_{col}, (1,1) for PWAV, and (1,1) for PWAT, and $a_1 = 0$ for T_{col}.

[17] Plots of the autocorrelation function (ACF) and partial autocorrelation (PACF) are usually examined to check the randomness and AR process parameter, p, of a time series. We applied these two tools before and after we removed ARMA(p, q) signals from the residuals. For all three time series, the original significant ACF and PACF values at all lags (not shown) were reduced to non-significant levels after the removal of ARMA(p, q) signals.

3.2.2. Generation of Monte Carlo Samples

[18] The variance of the estimated trend of a time series is not only a function of white noise variance, but also a function of the ARMA parameters. To evaluate we define $r = \max(p, q)$, and we arbitrarily set the first r elements of the time series as random numbers with variance equal to σ^2. The model then was spun up by computing the next 100 extra elements following equation (2). Another 321 elements were computed as one sample realization. A total of 100,000 realizations were generated for each of the variables - T_{col}, PWAV, and PWAT.

[19] The expectation for trend of every sample realization is zero, while its probability distribution function (PDF) is

Table 1. Estimated ARMA (p, q) Parameters

| Variable | Parameter |
	a_1	a_2	b_1	B_2	σ^2
T_{col}	0	0.784513	0.295777	−0.614191	0.112824
PWAV	0.771322	NA	−0.378879	NA	0.152011
PWAT	0.809517	NA	−0.437119	NA	0.138988

Gaussian. This is because every sample realization $\{y_t\}$ is generated by a random process and ARMA process, and its trend distribution is a linear combination of y_t. The probability density function (PDF) of the trends of North America T_{col}, PWAV, and PWAT are shown in Figure 2. The outcome of trend PDF is confirmed to be distributed normally according to the Shapiro-Wilk test [*Shapiro and Wilk*, 1965].

4. Results

[20] The values of T_{col}, PWAV, and PWAT trends, together with their standard deviations from the original OLS method, and the standard deviations for the trends from our simulations are shown in Table 2. The symbol $S_{\hat{\beta}_2}$ denotes the trend standard deviation derived from the original sample using the OLS method based on one single piece of time series for each variable. Its value, by the nature of unbiased estimation, should be very close to the mean of the trend standard deviations of the 100,000 time series realizations ($S_{\hat{\beta}_2}$), using the OLS method. However, $S_{\hat{\beta}_2}$ is far different from $S_{\hat{\beta}_2}^*$, the trend standard deviation, derived from the collection of 100,000 OLS trend estimates. With our new approach, the $t_{\hat{\beta}_2}^*$ for T_{col} is only ~37% of the original OLS t value ($t_{\hat{\beta}_2}$). This illustrates the importance of adopting the methodology described in this paper, without which the statistical significance of climate trends could be erroneously inflated by a conventional OLS approach (please also see the auxiliary material).[1]
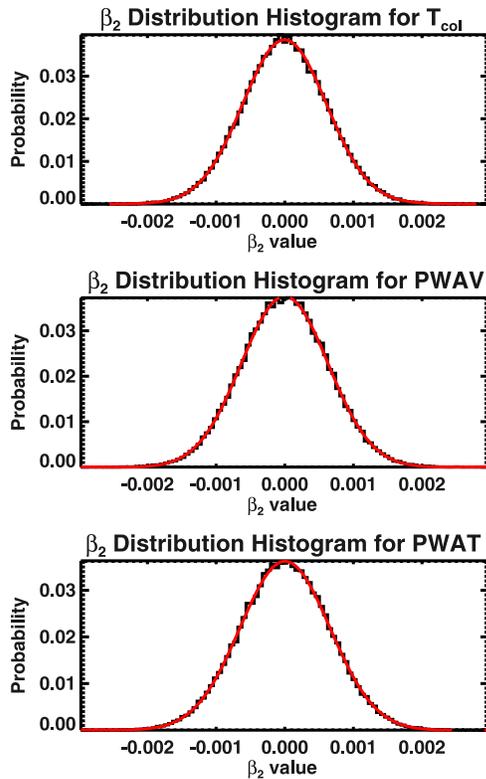
Figure 2. The trend distributions for (top) T_{col}, (middle) PWAV, and (bottom) PWAT. The black lines represent the experimental histograms, while the red curves denote Gaussian function.

[21] Setting the critical value of t as $t = 1.96$ (equivalent to a 95% confidence level), we found that the trend for T_{col} is significant, while the trends for PWAV and PWAT are far from being statistically significant. This result indicates that the water vapor response to lower-tropospheric temperature increase over North America is not as simple as presented by assuming a close relationship with the C-C equation or a constant relative humidity in a closed system.

[22] Given the 27-year mean of monthly T_{col} as 270.5K and the T_{col} warming trend of 2.06×10^{-3} K/month, a simple calculation will show that PWAV requires an increase of 4.53% if we assume constant relative humidity. However, PWAV only increases by 0.08% and PWAT only increases by 0.14% in the NARR data. In Figure 1a, the projected PWAV with constant relative humidity (black dash-dot lines) differs from the NARR PWAV (blue dashed lines) by a distinct rising angle.

[23] The discrepancy might be attributed to many reasons, such as moisture divergence, the change of surface evaporation and transpiration rate, etc. *Huntington* [2006] also reported that the constant relative humidity assumption could overestimate the warming-induced water vapor feedback. The replication of this lack of water vapor content over North America is a challenge for the modeling community to explain.

[24] Figure 1b shows the temporal variations of PWAV, PWAT, and T_{col} by season averaged over the domain. In this paper, we divide a year into four seasons following the convention: March, April, and May (MAM) for spring; June, July, and August (JJA) for summer; September, October, and November (SON) for autumn; December, January, and February (DJF) for winter. Note that we did not perform any simulation or statistical tests for the seasonal time series. It is obvious, however, that PWAV and PWAT trends in North America region are *inconsistent* with the trends projected by rising temperature if a constant relative humidity is assumed, which can be seen clearly by contrasting the brown and blue dashed lines with the black dashed line. In winter and spring, PWAV and PWAT trends actually have negative signs.

5. Discussion and Conclusion

[25] To detect the trend in any geophysical time series, the memory needs to be considered when performing the Student's t test. Unlike the GLS method, which considers the covariance structure across the time series, the ARMA takes into account the most recent memory of the time series itself (AR) and the most recent memory of random deviation (MA). In this research, we first assumed three time series (T_{col}, PWAV, and PWAT) as ARMA processes, after removing seasonality and the OLS trend. The Monte Carlo experiments were then conducted to generate 100,000 sample realizations with the same ARMA parameters for the T_{col}, PWAV, and PWAT time series. Regardless of the nature of the ARMA process, these realizations would create a normal distribution for the trend. The variance of trend can then be used to test the null hypothesis of the time series that there is a zero trend. In our cases, the variance estimated from our new approach can be as much as 2.7 times that of the original OLS variance.

[26] We found that for the domain we evaluated in the NARR, temperatures significantly increased (0.248 ± 0.0742 K/decade) according to the 27-year monthly data, but the precipitable water vapor (0.00619 ± 0.0755 Kg/m^2/decade) and total precipitable water (0.0108 ± 0.0782 Kg/m^2/decade) did not. This implies that a) PWAV and PWAT depend on net horizontal flux divergence

Table 2. OLS Trend ($\hat{\beta}_2$), OLS Trend Standard Deviation From the Sample ($S_{\hat{\beta}_2}$), Mean of OLS Trend Standard Deviations From the Simulation (S_{β_2}), OLS-ARMA Trend Standard Deviation From the Simulation ($S_{\beta_2}^*$), and Their Student's t Values[a]

	$\hat{\beta}_2$	$S_{\hat{\beta}_2}$	S_{β_2}	$S_{\beta_2}^*$	$t_{\hat{\beta}_2}$	$t_{\hat{\beta}_2}^*$
Tcol	2.06×10^{-3}	2.30×10^{-4}	2.26×10^{-4}	6.17×10^{-4}	9.13	3.34
PWAV	5.16×10^{-5}	2.77×10^{-4}	2.72×10^{-4}	6.28×10^{-4}	0.19	0.08
PWAT	8.99×10^{-5}	2.67×10^{-4}	2.61×10^{-4}	6.52×10^{-4}	0.34	0.14

[a]The units for $\hat{\beta}_2$ and its standard deviation are K/month for T_{col} and kg/m^2 month for PWAV and PWAT, respectively.

and surface water budget (evapotranspiration minus precipitation) as much as the implication of the C-C equation with a fixed relative humidity, and/or b) the atmosphere does not hold as much water as would occur with a fixed relative humidity. Our study domain is mostly over land, and this could contribute to the discrepancy in relative humidity between our findings and the global mean value.

[27] Our approach indicates that the multi-decadal trends in water vapor content on the regional scale are not yet well understood; a conclusion was also reached by *Spencer et al.* [2007]. We thus urge further evaluations of lower-tropospheric temperature and water vapor trends for other regions of the globe using the outlined statistical analysis methodology, as well as more examination of cloud-precipitation feedbacks in global and regional climate models.

References

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *2nd International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–281, Akad. Kiado, Budapest.

Arakawa, A. (2004), The cumulus parameterization problem: Past, present, and future, *J. Clim.*, *17*(13), 2493–2525.

Box, G. E. P., and G. M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco, Calif.

Brockwell, P. J., and R. A. Davis (2000), *Introduction to Time Series and Forecasting*, Springer, New York.

Chase, T. N., R. A. Pielke Sr., J. A. Knaff, T. G. F. Kittel, and J. L. Eastman (2000), A comparison of regional trends in 1979–1997 depth-averaged tropospheric temperatures, *Int. J. Climatol.*, *20*, 503–518.

Chambers, J. M., W. S. Cleveland, B. Kleiner, and P. A. Tukey (1983), *Graphical Methods for Data Analysis*, Wadsworth, Belmont, Calif.

Chelliah, M., and C. F. Ropelewski (2000), Reanalysis-based tropospheric temperature estimates: Uncertainties in the context of global climate change detection, *J. Clim.*, *13*, 3187–3205.

Core and Extended Writing Team (2007), *Intergovernmental Panel on Climate Change Fourth Assessment 2007*, Cambridge Univ. Press, Valencia, Spain.

Fomby, T. B., and T. J. Vogelsang (2002), The application of size-robust trend statistics to global-warming temperature series, *J. Clim.*, *15*, 117–123.

Hannan, E. J., and R. Rissanen (1982), Recursive estimation of ARMA order, *Biometrika*, *69*, 81–94.

Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Clim.*, *19*(21), 5686–5699.

Huntington, T. G. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, *319*, 83–95.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360.

Pielke, R. A., Sr., J. Eastman, T. N. Chase, J. Knaff, and T. G. F. Kittel (1998), 1973–1996 trends in depth-averaged tropospheric temperature, *J. Geophys. Res.*, *103*(D14), 16,927–16,933.

Shapiro, S. S., and M. B. Wilk (1965), An analysis of variance test for normality (complete samples), *Biometrika*, *52*(3–4), 591–611.

Soden, B. J., and I. M. Held (2006), An assessment of climate feedbacks in coupled ocean-atmosphere models, *J. Clim.*, *19*(14), 3354–3360.

Spencer, R. W., W. D. Braswell, J. R. Christy, and J. Hnilo (2007), Cloud and radiation budget changes associated with tropical intraseasonal oscillations, *Geophys. Res. Lett.*, *34*, L15707, doi:10.1029/2007GL029698.

Weatherhead, E. C., et al. (1998), Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.*, *103*(D14), 17,149–17,161.

Woodward, W. A., and H. L. Gray (1993), Global warming and the problem of testing for trend in time series data, *J. Clim.*, *6*, 953–962.

Zastawny, A. (2006), Calculation of solar and thermal radiation absorption in the atmosphere, based on the HITRAN data, *Meteorol. Atmos. Phys.*, *92*, 153–159.

————————————

   J. C. Lin, Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

   T. Matsui, Code 613.1, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

   R. A. Pielke Sr., Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA.

   J.-W. Wang, Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO 80309, USA. (jihwang@colorado.edu)

   K. Wang, Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA.