NOTES AND CORRESPONDENCE

On the Estimation of Trends in Annual Rainfall Using Paired Gauge Observations

NATHAN PALDOR

Department of Atmospheric Sciences, Institute of Earth Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

(Manuscript received 24 January 2007, in final form 19 December 2007)

ABSTRACT

A method was recently proposed for evaluating the impact of a perturbation, such as air pollution or urbanization, on the precipitation at a location by calculating the ratio between the precipitation at the perturbed location and that at a location believed to be unperturbed. However, this method may be inappropriate because of the high degree of variability of precipitation at each of the stations. To explore the validity of this approach, noisy annual rainfall records are generated numerically in an upwind, unperturbed station and in a downwind, perturbed station, and the time series of ratio between the annual rainfalls in the two stations is analyzed. The noisy rainfall records are 50 yr long, and the imposed trend for the downwind, perturbed station is

$$\frac{1}{100} \text{mm yr}^{-1}$$

while at the upwind station the variations in annual rainfall are purely noisy. Many pairs of noisy rainfall records are numerically generated (each pair constitutes an experiment), and in every experiment the slope of the linear best fit to the rainfall ratio yields an estimate of the trend of rainfall at the perturbed station. In the absence of noise, the trend of the rainfall ratio is explicitly related to the trend of rainfall at the perturbed station, but the natural rainfall variation at the stations completely masks this explicit relationship. The results show that in some experiments the trend line of the rainfall ratio has the opposite sign to the imposed trend and that in only about one-half of the experiments does the ratio’s trend line lie within \( \pm 75\% \) of the imposed trend. Trend estimates within \( \pm 25\% \) of the imposed trend are obtained in less than one-quarter of the experiments. This result casts doubt on the generality and validity of using trends of rainfall ratio between two stations to estimate trends of precipitation in one of these stations.

1. Introduction

Annual rainfall records at a single station have large variations that are sometimes larger than the mean annual rainfall so that the coefficient of variation exceeds 100\% (e.g., Sumner 1988). Even over continental areas where data from many (i.e., several hundred) stations are averaged, the variations in adjacent 30-yr windows can be as high as 25\% (Hulme and New 1997). Other studies have shown that high variability of annual rainfall records exists regardless of location or the exact length of the time series (Domroes et al. 1998). In many meteorological stations worldwide, the maximum annual rainfall in a 50-yr period is at least 3 times the minimal rainfall in that period, which amounts to variation of 50\% relative to the mean rainfall. (Note that this 50\% variation does not mean that the coefficient of variation is 0.5; it only implies that the span of values between the observed maximal and minimal rainfall is equal to the mean rainfall in the record.) Although the mean and standard deviation can be calculated for rainfall records, the meaning of these calculations is unclear because rainfall distribution is not Gaussian, as is clearly evident from its asymmetry: years with 50\% more rainfall than the mean occur more frequently than years with 50\% less rainfall than the mean. Sumner (1988, chapter 8) lists many subtleties associated with the distribution of annual rainfall, including the fact that the 30-yr averages vary significantly from one period to another.

The large variability, asymmetry, and nonnormal distribution of annual rainfall at a single station over several decades completely mask any long-term trend that might exist in that record. In an attempt to overcome this inherent large variability of annual rainfall at a
single station, the station-to-station rainfall ratio (i.e.,
the record of ratios between annual rainfall at one sta-
tion and that at another station in the same year) was
employed in recent studies to detect long-term trends in
one of these stations. The calculation of trend line of
station-to-station rainfall ratio for estimating the trend
of rainfall in one of the stations (presumably to circum-
vent the inherent large variation of annual rainfall at a
single station) has been used to underscore such effects
as air pollution there (Givati and Rosenfeld 2004, their
Figs. 2 and 3) and to quantify the impact that increasing
the concentration of condensation nuclei has on the
annual rainfall at one of these stations (Rosenfeld and
Givati 2006). The analysis of rainfall ratio reported by
Alpert et al. (2008) in many pairs of stations in Israel
has shown trends that are opposite to those calculated
by Rosenfeld and Givati (2006) and led to the opposite
conclusion regarding the effect of urbanization on the
precipitation downwind.

Elementary statistical reasons suggest that the vari-
ability of rainfall ratio is quantitatively different from
(and in general larger than) that of rainfall at a single
station. This can be easily demonstrated for two nor-
mally distributed rainfall records, X and Y (which is not
true for observed records). The station-to-station rain-
fall ratio \( r = X/Y \) is not normally distributed, because
division (and multiplication) is a nonlinear operation
while the normal distribution is invariant only under
linear operations (e.g., Feller 1966, p. 77). This consider-
ation motivates an examination of the relationship
between the trend line of station-to-station rainfall ra-
tio and the corresponding trend of rainfall at one of
these stations.

To accomplish this examination, the numerically gen-
erated rainfall record at one station is perturbed by
imposing a known, linear trend on it and the trend line
of rainfall ratio with another, trend-free (i.e., unpertur-
bled) station is calculated. This approach enables an
analysis of the relationship between the imposed trend
at a station and the calculated trend line of rainfall ratio.
The statistical characteristics of the numerically gen-
erated records used in this study mimic those of the
observed rainfall records (Givati and Rosenfeld 2004;
Rosenfeld and Givati 2006; Alpert et al. 2008) but the
year-to-year variability is smaller. Our findings demon-
strate that because of the large variation in the records
at the two stations the resulting trend line of rainfall ratio
is statistically insignificant and does not yield an
estimate of the trend in the perturbed station.

2. Numerically generated noisy rainfall records

The distribution of annual rainfall at a single station
has never been quantified, and even for daily rainfall
for which the Weibull distribution sometimes approxi-
mates the distribution (e.g., Barcelona; Burgueño et al.
2004), no universal distribution exists (Ananthakrish-
nan and Soman 1989). In the absence of a universal
distribution, the variation of annual rainfall in the two
stations of the current study was modeled simply as
white noise, which can be easily parameterized to gen-
erate a desired correlation between the rainfall records
in the two stations and enables reproduction of the re-
results reported here using a random number generator
only.

The 50-yr rainfall records in stations A (upwind) and
B (downwind) were generated numerically as follows:
The rainfall records in stations A and B, \( r(A) \) and \( r(B) \),
had initial noise-free values of 500 and 1200 mm yr\(^{-1}\),
respectively. These mean rainfall values are typical to
reported comparisons between perturbed and unpertur-
bled stations, as is the initial “enhancement factor” of
1200/500 = 2.4. The rainfall in B was decreased by 2
mm yr\(^{-1}\) every year so that by the 50th year the rainfall
there was 1100 mm yr\(^{-1}\). The resulting change in sta-
tion-to-station rainfall ratio, \( r(B)/r(A) \), associated with
this imposed rainfall trend at B is

\[
\frac{r(B)}{r(A)} \bigg|_{r=50} - \frac{r(B)}{r(A)} \bigg|_{r=1} = \frac{r(B)|_{r=50} - r(B)|_{r=1}}{r(A)} = \frac{(1200 - 2 \times 50) - 1200}{500} = -0.2.
\]

The annual variation in rainfall at stations A and B was
modeled as white noise that was added to the above
means (and the trend at B). As explained above, any
other distribution can be used, provided that the result-
ing rainfall records have the observed correlation of
about 0.8 for noise amplitudes that yield year-to-year
variation of rainfall similar to the observed variations.
At station A, the noise amplitude (i.e., the absolute
value of the maximum deviation of rainfall in any year
from 500 mm) was chosen as 250 mm yr\(^{-1}\) to insure a
low coefficient of variation (or standard deviation; see
below). At station B, the choice of noise level in the
annual rainfall is dictated by requiring that the rainfall
fluctuations in the two stations have a correlation of
about 0.8, which is the observed order of reported cor-
relation between pairs of stations (it guarantees that the
fluctuations in the two stations are not independent of
each other). Thus, the annual rainfall at station B was
set as the sum of four terms: 1) the rainfall at station A
(including noise), 2) an excess rainfall of 700 mm yr\(^{-1}\),
3) an additional noise of amplitude 175 mm yr\(^{-1}\), and 4)
an imposed trend of −2 mm yr\(^{-2}\) as discussed above.
where A and B had lower than the fluctuations reported in observed rainfall. The values used in Eq. (2) yield typical (not extreme) results. The standard deviations of the rainfall records generated by Eq. (2) in the experiments described below were found to vary between 125 and 165 mm yr⁻¹ at station A and 150 and 210 mm yr⁻¹ at station B. These year-to-year fluctuations in rainfall are considerably lower than the fluctuations reported in observed rainfall records. In addition, the rainfall records in stations A and B had $R^2$ (where $R$ is correlation coefficient) values that varied between 0.75 and 0.9, which is the same range as in pairs of upwind/downwind stations reported in the literature.

Other values of mean rainfalls, noise levels, and trends were used in Eq. (2), and they yielded results that are qualitatively similar to those reported below. The values used in Eq. (2) yield typical (not extreme) results.

A typical scenario of the numerically generated annual rainfall records in the two stations that follows from Eq. (2) is shown in Fig. 1, which underscores the small spread of annual rainfall in each station and the high correlation between the rainfall records in these stations ($R^2 = 0.86$). As expected, the trend line passes close to the mean values of 500 and 1150 mm yr⁻¹.

The next section addresses the confidence with which one can detect the imposed trend of −2 mm yr⁻¹ at station B (or its rainfall ratio equivalent of −0.2; see Eq. (1)] from the trend line of a 50-yr record of $\gamma_i = r(B)_i/r(A)_i$.

3. The trend line of B-to-A rainfall ratio and the trend at B

We now attempt to reconstruct the imposed trend in station B from the time series of rainfall ratio. The method described below is an adaptation of that used in the observational studies cited in the introduction.

The trend line of the station-to-station rainfall ratio $\gamma_i = r(B)_i/r(A)_i$ yields a linear (least squares) approximation $\Gamma_i$, and the difference between its initial ($i = 1$) and final ($i = 50$) values, $\Gamma_{50} - \Gamma_1$, is the net change in rainfall ratio over 50 yr. According to Eq. (1), without noise and for the imposed rainfall trend in B of −2 mm yr⁻², $\Gamma_{50} - \Gamma_1$ should equal −0.2.

To increase the statistical significance, every run reported here represents the results of 12 numerical experiments (increasing the number of experiments in a run to more than 12 had little effect on the results). In each experiment, different sets of random numbers $\xi_i$ and $\zeta_i$ were used in Eq. (2) to generate the two rainfall records similar to those shown in Fig. 1. Two such experiments from a single run along with the corresponding trend lines are shown in Fig. 2: In the upper panel, the trend line of rainfall ratio increases from 2.205 in the first year ($i = 1$) to 2.560 in the 50th year ($i = 50$) so that the net change in $\Gamma_i$ during the 50 yr is +0.36. In the lower panel, the trend line of rainfall ratio decreases from 2.851 in the first year ($i = 1$) to 2.106 in the 50th year ($i = 50$) so that the net change in $\Gamma_i$ during the 50 yr is −0.75. These two estimates differ from the prescribed trend of rainfall ratio (imposed as the −2i term in the rainfall at station B) of −0.2 by 300%!

The two panels of Fig. 2 represent only 2 (out of 12) experiments in which the change in the trend line differed significantly from the expected change in rainfall ratio given the imposed trend at B. A more general assessment of the usefulness of calculating trend lines of noisy rainfall ratios is achieved by calculating the net change in the trend line of rainfall ratio in 20 independent runs (where each run consists of 12 experiments). An experiment is considered to be accurate if the net change in the trend line of rainfall ratio, $\Gamma_{50} - \Gamma_1$, lies near −0.2. The difference between $\Gamma_{50} - \Gamma_1$ and −0.2 is quantified by the relative deviation of $\Gamma_{50} - \Gamma_1$ from −0.2; that is, $[(\Gamma_{50} - \Gamma_1) - (−0.2)]/−0.2 = 1 - 5(\Gamma_{50} - \Gamma_1)$. The deviation levels used here are 25%, 50%, and 75%, and they correspond to $\Gamma_{50} - \Gamma_1$ values in the ranges from −0.15 to −0.25, from −0.1 to −0.3, and from −0.05 to −0.35, respectively. The number of experiments in which $\Gamma_{50} - \Gamma_1$ lies inside a given level of deviation varies among runs. However, the average of these numbers over 20 runs is very robust and changes only slightly (a small fraction of 1.0) from one
set of 20 runs to another. The results of the calculations of the average over 20 runs (for each range of accuracy) are shown in Fig. 3, which clearly demonstrates that, even at the 75% accuracy range, only about one-half of the runs resulted in accurate estimates, whereas at the 25% level slightly less than one-quarter (3 of 12) of the experiments resulted in accurate estimates.

4. Discussion

Observed annual rainfall records have year-to-year variations that are larger than the variations simulated in this study. In this study, the standard deviation of annual rainfall at station A (where the mean rainfall is 500 mm yr\(^{-1}\)) varies between 125 and 165 mm yr\(^{-1}\), and at station B (where the mean rainfall is 1150 mm yr\(^{-1}\)) it varies between 150 and 210 mm yr\(^{-1}\). The statistical characteristics of the natural rainfall variation at a single station vary drastically between stations, and in general the global map of the coefficient of variation (defined as the standard deviation divided by the mean) of annual rainfall reveals that the variation decreases with the rainfall [cf. Figs. 3.4 and 3.15 in Riehl (1979)]. Israel and northern California [the sites where the studies of Givati and Rosenfeld (2004), Rosenfeld and Givati (2006), and Alpert et al. (2008) were conducted] lie within the “over 40%” and “30%–40%” contours of Fig. 3.15 in Riehl (1979). The coefficient of variation in the study described in this paper varies between 0.1 (150/1150, at station B) and 0.3 (165/500, at station A), and so the simple white-noise distribution assumed in Eq. (2) (intended merely to reproduce a typical observed distribution of the asymmetric annual rainfall) yields means and standard deviations that fall within the observed ranges. Our results highlight the uncertainty in the estimation of trends in the downwind station based on the calculation of the rainfall ratio between this and an upwind unperturbed (i.e., trend free) station.

The results presented here should not be taken to imply that the microphysical arguments of Rosenfeld and Givati (2006) and Givati and Rosenfeld (2004) regarding the suppression of rainfall by the increase of air pollution are incorrect—they only imply that the observational evidence lacks statistical significance. The results reported here demonstrate that even when the known trend in the downwind station is of one sign, the large natural year-to-year variations (modeled here as white noise) can reverse the sign of the trend line of rainfall ratio with an upwind station. In fact, by evaluating the rainfall ratio between many pairs of downwind and upwind stations in Israel, Alpert et al. (2008) have demonstrated that in 50% of the pairs the change in the trend line of rainfall ratio during the 1954–2004
period is positive, which contradicts the results of Givati and Rosenfeld (2004). The fact that these two studies yield opposing estimates of the trends in pairs of stations in Israel is consistent with the conclusion of the present study that the evaluation of trends in rainfall at a station based on the trend line of rainfall ratio lacks statistical validity. One only needs to regard the year-to-year variation in the observed records as the counterparts of the white noise in this study to realize that the two observationally based estimates are possible.

The relation between the trend of the annual rainfall at a single station and the trend line of rainfall ratio between this station and another station is masked by the high variability of the records in each of these stations and by the low statistical significance of the trend line calculation.

Acknowledgments. The study was supported by ISF Grant 579/05 to The Hebrew University of Jerusalem.

REFERENCES


