On the Detection of Statistical Heterogeneity in Rain Measurements

A. R. JAMESON

RJH Scientific, Inc., El Cajon, California

M. L. LARSEN[a]

Department of Physics and Astronomy, College of Charleston, Charleston, South Carolina

A. B. KOSTINSKI

Department of Physics, Michigan Technological University, Houghton, Michigan

(Manuscript received 14 September 2017, in final form 23 April 2018)

ABSTRACT

The application of the Wiener–Khintchine theorem for translating a readily measured correlation function into the variance spectrum, important for scale analyses and for scaling transformations of data, requires that the data be wide-sense homogeneous (stationary), that is, that the first and second moments of the probability distribution of the variable are the same at all times (stationarity) or at all locations (homogeneity) over the entire observed domain. This work provides a heuristic method independent of statistical models for evaluating whether a set of data in rain is wide-sense stationary (WSS). The alternative, statistical heterogeneity, requires 1) that there be no single global mean value and/or 2) that the variance of the variable changes in the domain. Here, the number of global mean values is estimated using a Bayesian inversion approach, while changes in the variance are determined using record counting techniques. An index of statistical heterogeneity (IXH) is proposed for rain such that as its value approaches zero, the more likely the data are wide-sense stationary and the more acceptable is the use of the Wiener–Khintchine theorem. Numerical experiments as well as several examples in real rain demonstrate the potential of IXH to identify statistical homogeneity, heterogeneity, and statistical mixtures. In particular, the examples demonstrate that visual inspections of data alone are insufficient for determining whether they are wide-sense stationary. Furthermore, in this small data collection, statistical heterogeneity was associated with convective rain, while statistical homogeneity appeared in more stratiform or mixed rain events. These tentative associations, however, need further substantiation.

1. Introduction

There is strong interest in understanding the temporal and spatial scales of rainfall. Observations and numerical models produce results on many different scales, extending from that of a single rain gauge up to scales of one to hundreds of kilometers from forecast models. A significant challenge, then, is to retain fidelity when translating among all of these different scales. In particular, more realistic forecast model results often depend upon integrating observations into the model, while the utility of numerical forecasts often depends upon the translation of model output down to scales of, say, urban flooding.

One of the most powerful tools for investigating various scales in rain is, of course, the power spectrum (e.g., Crane 1990; Kiely and Ivanova 1999), the Fourier transform of the relatively easily observed autocorrelation function (via the Wiener–Khintchine theorem; Wiener 1930; Khintchine 1934). The latter often provides the most convenient route to the power spectrum from observations, while filtered power spectra can yield autocorrelation functions useful for interpreting measurements collected over finite temporal or spatial domains of different sizes.

However, the validity of the Fourier transform relation between the autocorrelation function and the variance spectrum depends upon the data being statistically stationary

Denotes content that is immediately available upon publication as open access.

[a] Additional temporary affiliation: Michigan Technological University, Houghton, Michigan.

Corresponding author: A. R. Jameson, arjatrjhsci@verizon.net

DOI: 10.1175/JTECH-D-17-0161.1

in time or homogeneous in space (Jerison et al. 1997). Stationarity (homogeneity) can have a much more demanding strict sense meaning and a weaker wide-sense meaning. In the strict sense, all moments of the probability distribution of a variable are the same everywhere or at all times over a domain in its entirety, while in the wide-sense stationarity (WSS), it is sufficient to have a constant global mean and constant variance everywhere or at all times over all the samples. What is meant by "mean value?" Over a domain there are what can be called "global mean" values applicable over the entire domain. At any particular location, however, there are local mean (LM) values, which represent the spatially and temporally correlated statistical mixtures of these global means. The LMs do not represent real changes in the global mean values, however.

Obviously, then, the definition of WSS is necessarily scale dependent, since all sets of samples are finite in size so that when these terms are used, it is implicitly understood that they apply from the smallest resolved scales up to the largest dimension of the sample domain even though correlations can introduce local variability into the LMs, for example. To rephrase slightly, WSS can apply when there are no temporal or spatial variations or gradients of the global means or changes in the variance over the largest dimensions. This does not mean, however, that gradients in the set of observed LMs cannot exist depending upon the particular expression of the correlation functions.

In natural rain it is highly unlikely that the much more demanding strict sense stationarity (homogeneity) ever occurs (Nason 2006, p. 4; Schleiss et al. 2014; Serinaldi et al. 2018). The complex four-dimensional structure of rain will often produce statistical mixtures exhibiting correlations. Fortunately, the Wiener–Khintchine theorem requires only WSS. However, in rain, one cannot simply look at a set of data and tell whether it is WSS, as we will clearly demonstrate later. Yet, in almost all studies of rain, these conditions are assumed to be true in no small part because of the lack of a method to test the data.

The determination of whether data are WSS is not a trivial problem (Nason 2006, p. 2.):

> A tricky question is how can you know whether a time series is stationary or not? There are various tests for stationarity. As well as suffering from all of the usual issues of statistical testing . . . tests of stationarity tend to test against particular alternative models of particular types of non-stationarity. For example, test A might well be powerful at picking up non-stationarities of type A, but it has no power at detecting those of type B.

Thus, most approaches toward determining stationarity (homogeneity) that are found in the literature depend upon developing a statistical model of the data. As stated above, sometimes they work and sometimes they do not, and they all make assumptions about the data. Hence, in this work we focus on using the data itself without any assumptions or potentially misleading models. This is done by looking directly for changes in the global mean values and the variance. When found, one then knows whether to attempt to remove these effects so that the Wiener–Khintchine theorem can be applied safely. For example, Schleiss et al. (2014) develop one such approach for some circumstances.

The purpose of this work, then, is to suggest a method and an index for evaluating the temporal statistical stationarity or the spatial statistical homogeneity of rain observations in order to detect when data are or are not appropriate for the Wiener–Khintchine transform between the correlation function and the power spectrum without further processing of the data. This will be accomplished by looking at a measure of changes in the variance of the observed variable and by looking for any evidence of multiple global mean values that could indicate shifting global means. (Note that in this work, while technically sloppy, *statistical stationarity* and *statistical homogeneity* are used interchangeably for convenience, noting that when we use one of the terms, we are really talking about both simultaneously. We will use WSS in this sense throughout the remainder of the article.) However, before looking at real rain, in the next section a background discussion is provided to clarify the subsequent analyses of both simulated data and several real observations in rain.

2. Background

As just discussed, there are, then, two components to WSS that have to be evaluated, namely, the global mean and the variance. Recently, a straightforward method for detecting changes in the variance of a variable were developed in Anderson and Kostinski (2010, 2011). In that approach successive record highs and record lows are counted in both the forward and backward directions. When the data are WSS, the total counts (T) in each direction will be nearly the same so that $\alpha = T_{\text{forward}} - T_{\text{backward}}$ approaches zero. They show that deviations from the null depend upon two factors, namely, 1) the statistical fluctuations (measured by the standard deviation, σ_α) which, in turn, are functions of the length of the time series (the total number of observations); and 2) the success in removing any trends in the LMs that can distort α. It is important, therefore, to compute the LM curve and to subtract this curve from the raw data in order to get the fluctuations unbiased by any underlying structure of correlated LMs or systematic changes in the global mean values when there is more than one such value. We assume, therefore, that the curve can be divided into a component associated with changes in the

LMs and/or global means and a fluctuating component. We then estimate this mean curve by computing a least squares weighted spline fit over a distance sufficient to yield a zero mean distribution of fluctuations (usually on the order of 1–2 times the observed decorrelation length). The record statistics are then determined for the fluctuating component calculated as the difference between the observed and the mean curve.

This assertion was tested using numerical experiments by constructing several different random time series for Gaussian, gamma, uniform, and exponential distributions of a random variable each having a different but constant mean. These series were then exponentially correlated using the copula technique using different times to decorrelate for each. The comparison of the uncorrelated and correlated times series yielded the same α. Moreover, this was true regardless of the underlying distribution of the random variable as pointed out by Anderson and Kostinski (2010). This is important because the distribution for α often appears to be Gaussian with zero mean but having a variance dependent on the sample size as discussed further below. Therefore, throughout the remainder of this work, when we refer to α, we mean the α for this fluctuating component of the observations. We will also assume that the distribution of α will be Gaussian with zero mean when testing whether the data are WSS.

Furthermore, Anderson and Kostinski (2010, 2011) show that the frequency distribution of α is often well approximated by the normal distribution having zero mean when all the trends have been accounted for properly. Thus, we will consider deviations as being statistically significant only when they exceed $1.5\,\sigma_\alpha$. Thus, departures from constant α will be considered to be significant only when the relative dispersion $\mathrm{RD}_\alpha = |\alpha|/\sigma_\alpha > 1.5$. It is also important to point out here that this approach applies only to continuous variables, such as the rainfall rate or raindrop concentrations.

Anderson and Kostinski (2010, 2011) argue that when statistical stationarity prevails, σ_α depends on only the length of the observations regardless of the underlying frequency distribution of the data. We have verified this by simulating random samples drawn over a wide range of different frequency distributions of the data. Furthermore, for a series having N elements, Glick (1978) derived the variance for the number of records, say, in the forward direction F, namely,

$$\sigma_F^2 = 2\left(\sum_{i=1}^{N}\frac{1}{i} - \sum_{i=1}^{N}\frac{1}{i^2}\right). \tag{1}$$

The factor of 2 occurs because the records consist of new highs and new lows that are statistically independent. As Glick (1978) shows, when N is sufficiently large, the first term in the brackets goes to $\ln(N)$ + Euler's constant (0.5772...), while the second terms goes to $\pi^2/6$.

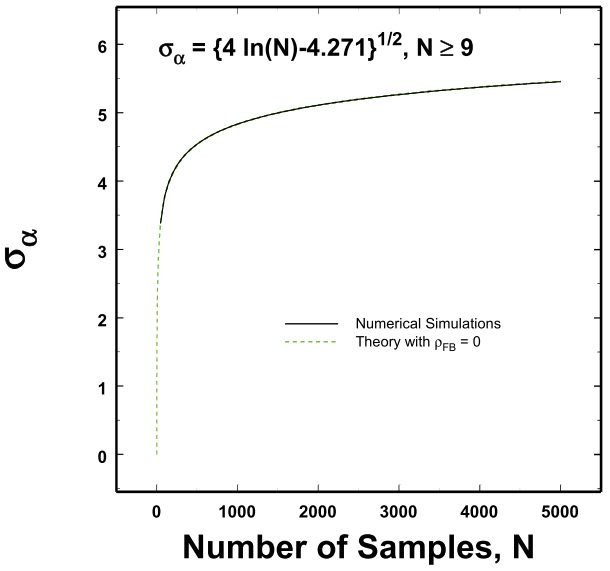

FIG. 1. The relation between the standard deviation σ_α of α to N observations determined from theory and confirmed using numerical simulations. The various lines are explained in the text, but the equation is used in the subsequent analyses.

Since α is the difference between the forward and backward B direction, it follows that

$$\sigma_\alpha^2 = \sigma_F^2 + \sigma_B^2 - 2\rho_{FB}\sigma_F\sigma_B, \tag{2}$$

where ρ_{FB} is the correlation coefficient between the standard deviations of the record counts in opposite directions. It is not obvious what ρ_{FB} should be. Instead, we use numerical simulations to derive σ_α as a function of the sample size N as plotted in Fig. 1. When we compare the numerical results with the theory, we find that at large N the two agree when $\rho_{FB} = 0$, so the forward and backward counts are statistically independent random variables. Expression (2) then becomes

$$\sigma_\alpha^2 = 2\sigma_{F,B}^2 = 2\left(\sum_{i=1}^{N}\frac{1}{i} - \sum_{i=1}^{N}\frac{1}{i^2}\right), \tag{3}$$

where the subscripts F, B refer to the forward or backward directions, respectively. While the summations are not all that difficult, it is often simpler to find a parametric fit that covers most realistic values of N. One such fit that is justified theoretically is shown in Fig. 1. For large N the leftmost sum in (3) approaches $4\ln(N) + 4\gamma$, where γ is the Euler constant 0.5772..., while the rightmost sum approaches a constant value $4\pi^2/6$ (see Glick 1978;

Anderson and Kostinski 2010). When all of these terms are combined, we have that

$$\sigma_\alpha = [4\ln(N) - 4.271]^{1/2}. \quad (4)$$

This theoretical equation matches the numerical simulations, deviating at the most by less than 2% for $N = 9$, so this expression is completely adequate for $N \geq 9$ as clearly demonstrated in Fig. 1.

As one might expect, σ_α grows with increasing N because there are then more opportunities for fluctuations. Hence, in these analyses, α is first calculated for the fluctuations from the LM curve. The absolute value of α is then divided by the expected σ_α computed using (4) to yield the relative dispersion RD_α. Values less than 1.5 are taken to be just a reflection of statistical random fluctuations, while values greater than 1.5 are taken to indicate values 87% likely to be real and not as a result of statistical fluctuations. Of course, one is free to select a different threshold and, obviously, when $N_c > 1$ (number of global means) this contribution usually becomes of more academic interest, since we already know, then, that the data cannot be WSS except for statistical mixtures.

As mentioned, the other component of WSS to consider is the global mean value. It would seem almost trivial to fit a ''mean'' curve to a plot of observations to determine whether the global mean value is changing. This is a misconception because the local means are statistical fluctuations that are spatially and temporally correlated. Even data that are WSS can exhibit significant and systematic fluctuations about one global mean value just because random realizations can have spatial and temporal correlations. So, how do we detect the set and number of contributing global mean values?

In past work (Jameson 2007; Jameson and Heymsfield 2013, 2014; Jameson 2015) a Bayesian inversion method was developed for estimating the probability distribution of the global mean values of a series of observations. While details may be found in the references just mentioned, briefly, one considers a range of global mean values for an assumed particular form of a distribution (the so-called likelihood distribution). Each observation can then be associated with each of these global mean values to some degree of probability. As discussed in the abovementioned references, the form of the assumed distribution is not critical as long as it is physically reasonable and single peaked. Here we use a normal distribution. Over the entire set of observations, the probabilities at each mean value are then summed and normalized to unity to yield a final estimate of the probability distribution $P(C)$ of the mean values, C, themselves. The most likely component mean values N_c are those associated with local maxima in this distribution determined from first derivatives and second derivatives (see the appendix).

Sometimes, though, the components are not of sufficient magnitude to produce distinct maxima. In those cases, the combinations of components often produce inflections in the distribution that can, then, often be identified using second derivatives as illustrated in the appendix. The detection of these features depends on them being sufficiently separated with respect to the variance used in the Bayesian likelihood distribution. Since this variance is an independent parameter, it can be specified optimally to take into account the width of the distribution of the data and the total number of observations. This is discussed in further detail in the appendix. It is important to point out, however, that the results are not sensitive to correlations because they do not depend upon whether the observed values are clustered or spread out in time or space.

The whole idea, then, is that statistical heterogeneity can occur if there is more than one global mean value, that is, when $N_c > 1$. This test may fail, however, for statistical mixtures that are WSS. However, even in that event, at any one instant of data collection by sensors in a network, the local mean values will, of course, vary from location to location, so in that sense even then those observations may be considered statistically heterogeneous.

The remaining alternative is that $N_c = 0$; that is, there are no peaks in the distribution of the means. In that case every point is a new extremum in a set of monotonically increasing or decreasing values. Hence, the data are changing everywhere, so they cannot possibly be WSS. In such circumstances RD_α would also be large as pointed out by Anderson and Kostinski (2010).

In addition to these considerations of the global mean values, statistical heterogeneity can also result from the variance of the observable changing over the domain. Because of statistical fluctuations, we consider only $\alpha \geq 1.5\sigma_\alpha$ to be a real indicator of the presence of trends in the data. Table 1 summarizes the domains of statistical homogeneity and heterogeneity based upon the values for N_c and RD_α.

For convenience these two quantities can be potentially combined in many different ways into a single number or an index of statistical heterogeneity (IXH). One possible definition is

$$\mathrm{IXH} = \left[\frac{H\left(\frac{\mathrm{RD}_\alpha}{1.5} - 1\right)\left(\frac{\mathrm{RD}_\alpha}{1.5} - 1\right) + (N_c - 1)}{2}\right], \quad (5)$$

where H is the Heaviside unit step function, so RD_α contributes only when it exceeds 1.5.

TABLE 1. Domains of statistical heterogeneity and homogeneity defined by the values of N_c and RD_α. Statistical mixtures can be either homogeneous or heterogeneous depending on whether all the components have the same correlation function.

No. of components	$RD_\alpha \le 1.5$	$RD_\alpha > 1.5$
$N_c = 0$	Not WSSs	Not WSS
$N_c = 1$	WSSs	Not WSS
$N_c \ge 1$	If $RD_\alpha = 0$, then WSS/otherwise, not WSS	Not WSSs

This definition is not entirely arbitrary. Obviously for the data to be WSS, there can be only one global mean value; that is, N_c must equal unity unless the relative dispersion is zero (Table 1). Second, for the commonly observed normal distribution of α, we have specified arbitrarily that RD_α must exceed 1.5, so statistical fluctuations will exceed this value only about 13% of the time. Hence, IXH has that kind of statistical accuracy as well. This, of course, can be adjusted by the user.

Aside from the special case discussed above when $N_c = 0$, there are also a few special cases to consider as well, such as when N_c and RD_α are contradictory, that is, when

$$N_c = 1,\ RD_\alpha \ge 1.5, \tag{6a}$$

$$N_c \ge 1,\ RD_\alpha < 1.5. \tag{6b}$$

When $N_c = 1$ but $RD_\alpha > 1.5$, then the data are statistically heterogeneous solely because of the changes in the variance. On the other hand, when $RD_\alpha < 1.5$ while $N_c > 1$, there is, then, more than one contributing global mean value, so we can identify these data as likely being a statistical mixture.

This index in (5) is chosen, in part, because a value of IXH greater than 0.0 can then be taken as a measure of the statistical heterogeneity. IXH will be this small only when $N_c = 1$ and $RD_\alpha < 1.5$. As values of IXH increase, then, the more likely it is that the data are statistically heterogeneous as the statistical reliability of RD_α increases. In this work, therefore, we consider IXH > 0.0 to be a minimum indicator of statistical heterogeneity. Larger values suggest that the data should first be processed in some manner (e.g., the removal of any trends) before applying the Wiener–Khintchine theorem. In the next section we apply these discussions to real observations in rain.

Examples of the entire analysis using numerical simulations for both statistically heterogeneous and homogeneous cases are presented below (Figs. 11 and 12). However, there are other considerations as well. The observed N_c will depend upon how much each component contributes to the overall distribution $P(C)$ of global mean values C. As the contribution from one component decreases, it may well become invisible. At that point, however, initial calculations show that when they become less than 10%–20% of $P(C)$, their contributions to statistical heterogeneity become negligible as well. It is also possible that even with equal contributions from, say, two components, as the peaks associated with each move closer and closer together, one peak may well disappear into the other, once again becoming invisible. However, again, the contributions of the two are then essentially merged, so trying to distinguish homogeneity from heterogeneity based upon N_c becomes academic with likely little effect on whether one should use the Wiener–Khintchine theorem, depending upon the statistics of α. These possibilities should always be kept in mind, however. Thus, the approach used here, then, may technically miss some cases of statistical heterogeneity, but those that are detected are likely correct and those that are missed are probably not all that important with regard to the use of the Wiener–Khintchine theorem.

3. Observations in rain

a. *Single time series observations*

In this section we will look at 11 sets of rain observations. The first is of Joss–Waldvogel (JW) impact disdrometer data provided by the late Professor Carlton Ulbrich (Clemson University) of almost a 17-h rain observed at 1-min temporal resolution. The rainfall rates R were calculated for each of the 1-min samples (Fig. 2). Looking at the observations by themselves, the data appear to be statistically heterogeneous with a positive gradient ∇ (slope of a linear fit) over all the data. Indeed, α is found to be 44, so the RD_α is 9.10. Nevertheless, the Bayesian inversion yields only one component near $2\,\mathrm{mm\,h^{-1}}$. Consequently, IXH = 2.53, so the statistical heterogeneity is driven by the changes in the variance of the rainfall as discussed regarding (6a).

We next consider a rain event with an initial convective component followed by a showery period and finally a period of stratiform rain (Fig. 3) measured over a small network of disdrometers located near Charleston, South Carolina [for a complete description, see Jameson et al. (2015)]. Obviously there is considerable variability in the average rain rate, so we expect the rain to be statistically heterogeneous. The Bayesian inversion reveals several components to the distribution of the global mean values, so IXH = 7.42—a very large value. Hence, when considered in their entirely, these data are statistically heterogeneous (nonstationary). However, sometimes subdividing the data can be very useful.

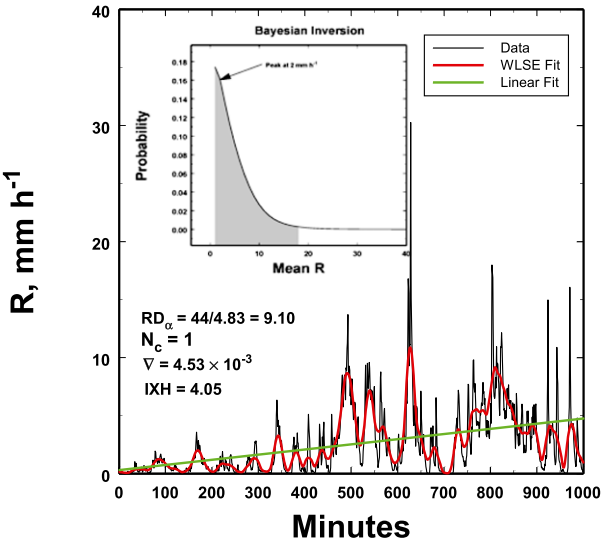

FIG. 2. Observations of rain in South Carolina. Despite there being only one Bayesian component, these data show significant statistical heterogeneity because of the increasing variance in time. The shading in the inset represents 99% of the distribution of the global mean values inferred from the Bayesian inversion. The term R has units of mm h^{-1}, and ∇ is the gradient of the linear fit across the data (green).

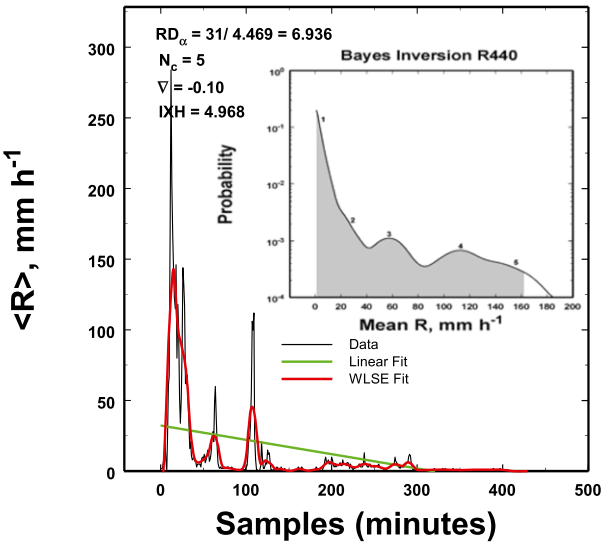

FIG. 3. The average values over a small network of 19 video disdrometers of a convective–stratiform rain event lasting 440 min as described further in the text. These data as a whole are statistically heterogeneous, but they can be divided into a statistically homogeneous section and a statistically heterogeneous section as presented next.

For example, we can divide the data in Fig. 3 into a convective period (0–115 min), a stratiform period (308–430 min), and a transition period that is a mixture of the two (190–308 min). As Fig. 4 shows for the convective data, much of the statistical heterogeneity of these data derives from the variability of the global mean values. The statistics were computed for each of the 19 detectors over the time period and then averaged to yield the average values indicated in the figure by the brackets <..>. For this component, then, the average IXH is computed to be 2.50—well within the domain of statistical heterogeneity. Since RD_α is not all that large, we conclude the heterogeneity of these data is being driven largely by the significant number of different components ($N_c = 6$) contributing to the rain.

In contrast, in Fig. 5 the stratiform component appears quite steady, having a very small overall gradient ∇. Just as for the convective rain, $RD_\alpha \cong 1$—a statistically marginal value. However, unlike the convective rain, there is now only one component in the Bayesian inversion. Consequently, the average IXH is 0.0, so these data are WSS.

If we then look at the transition or mixed region (Fig. 6), we might expect to find something in between these two domains. We do not. Even though apparent spikes in the rainfall rate appear presumably as small convective elements pass through the stratiform rain, there is only a weak overall negative gradient and there is still only one component in the distribution of the global mean values from the Bayesian inversion. Moreover, this time the influence of α is even smaller, since $RD_\alpha = 0.55$, so IXH = 0. Thus, in spite of spikes in the rainfall rate, the data are WSS, so one cannot simply tell necessarily by a visual inspection of the measurements whether statistical homogeneity prevails.

This realization is substantiated by another example (Fig. 7) of a mixture of bursts of heavier rain embedded

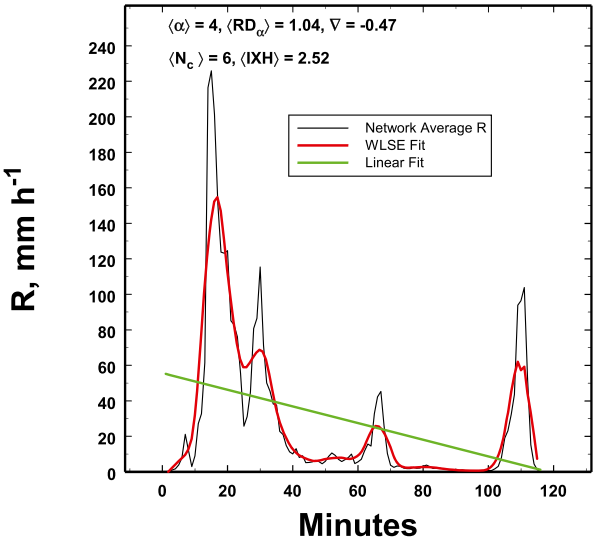

FIG. 4. The first approximately 120 min of the data in Fig. 3 for the convective rain in which the data are statistically heterogeneous.

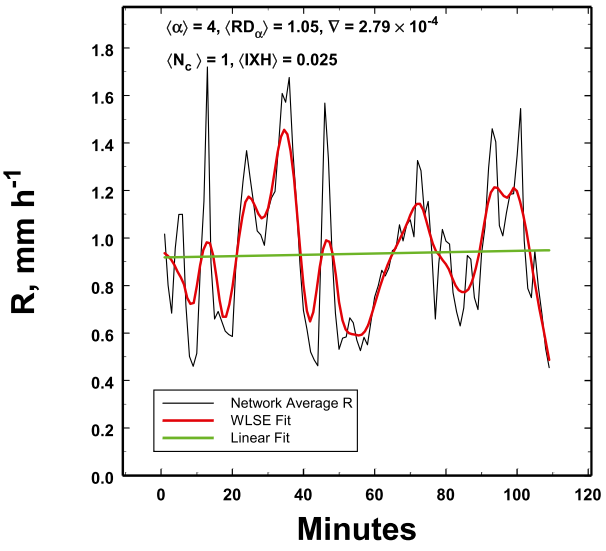

FIG. 5. Analysis of the stratiform period of rain during the 440-min rain event shows that it is statistically homogeneous as confirmed by the analyses of the global mean and fluctuating components as discussed in the text.

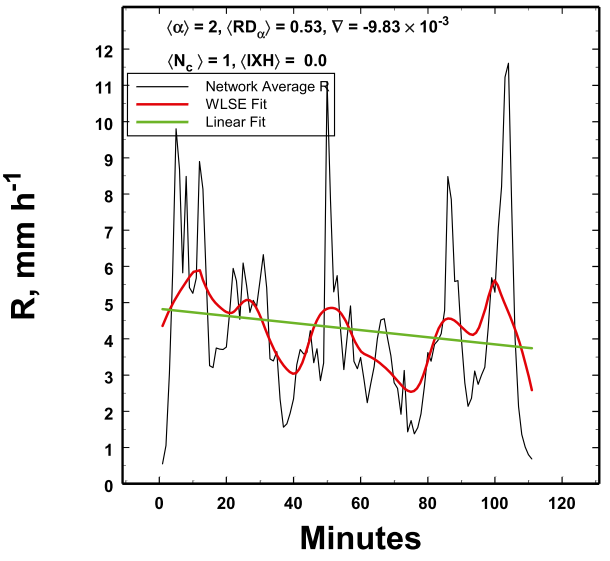

FIG. 6. The mixed period of rain (~200–300 min) in the 440-min rain event when background rain has occasional convective rain elements moving through, giving the appearance of being statistically heterogeneous. Despite this, the data are also statistically homogeneous as discussed in the text.

in a background of steadier rain collected over the network of disdrometers beginning at 1747 UTC 26 December 2013 during an 80-min period (previously described in Jameson et al. 2015). While the linear trend in the data (∇) remains small, it is not as tiny as for stratiform rain. Furthermore, there are three peaks in the Bayesian inversion, but this time the changes in the variance are also substantial with $RD_\alpha = 2.75$. Consequently, even though appearing quite similar to the data in Fig. 6, these data are significantly statistically heterogeneous with IXH = 1.88.

There are other examples as well from a different source, namely, JW disdrometer measurements from the National Aeronautics and Space Administration (NASA) rain facility at Wallops Island, Virginia, kindly provided by David Wolff. On 6 March 2013 a winter storm in the Ohio River valley made its way eastward toward the East Coast. Ahead of the storm center a warm front moved up the East Coast. A line of convection formed ahead of the warm front, and as it moved over Wallops Island, it produced two hours of heavier rain with another two hours of less intense rain. First, we consider the more intense rain in Fig. 8. In this case, the rain has a meaningful $RD_\alpha = 1.80$ as well as three main contributions to the distribution of the global mean values of R ($N_c = 3$). Consequently, these data are statistically heterogeneous (IXH = 1.10).

On the other hand, in the second 2-h period of lighter rain, there is one global mean value with correlated random fluctuations about that global mean that gives rise to a small ∇. These LM values (red line) are subtracted before computing α. The plot then reveals a clear trend in the variance. Indeed, $RD_\alpha = 2.70$, but in this case $N_c = 1$. The result is that IXH = 0.40, so these data are statistically heterogeneous because of changes in the variance. Hence, at all times one must consider both RD_α and N_c (Fig. 9), but they may be considered to be a statistical mixture because $RD\alpha = 0$.

b. *Multiple simultaneous time series observations*

While these examples above show the potential applicability of IXH, another factor that deserves consideration is the potential dependency of IXH on measurements by a single instrument; that is, rain is, of course, multidimensional, so a reasonable question is, Just how sensitive is IXH to a particular location and to the particular set of observations by one instrument?

This is a difficult question to address in general, However, we can get a feeling for potential sensitivity of IXH to spatial variations by returning once again to the observations from the disdrometer network already previously considered except that this time we can consider each instrument independently. In other words, we first perform this same kind of analysis on the temporal data from each instrument and then combine them spatially through interpolation using commercially available software and the conservative interpolation scheme of Sibson (1981) and Watson (1992). This interpolation does not create artificial features, such as artificial spatial maxima or minima, yet it precisely retains all the observations.

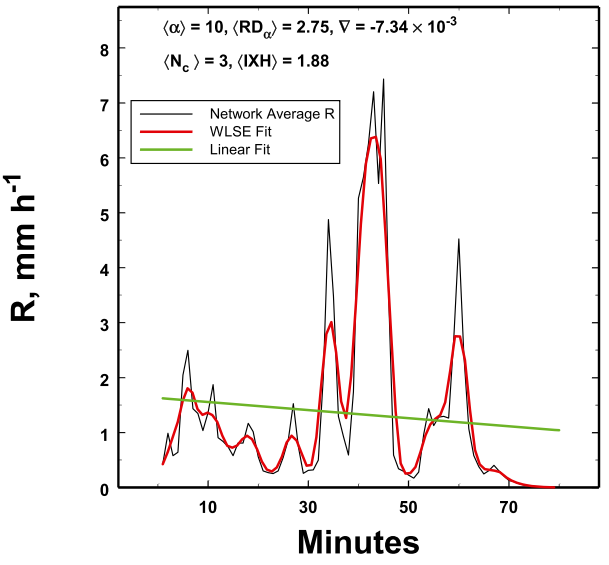

FIG. 7. As in Fig. 6, this is an 80-min mixture of different rain intensities on 26 Dec 2013, but unlike Fig. 6, this rain is statistically heterogeneous.

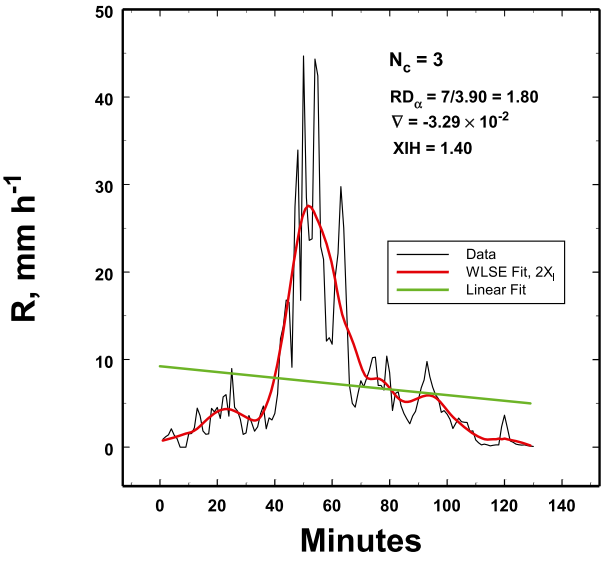

FIG. 8. Data from the NASA Wallops Island rain facility through a convective rainband passage preceding a warm front. These data are clearly statistically heterogeneous as discussed in the text.

We begin with the convective rain as illustrated in Fig. 10a. As expected from Fig. 4, strong statistical heterogeneity extends across most of the entire area. A striking feature, though, is the apparent variability of IXH even across the 100 m × 72 m area, showing the danger of considering using results from one instrument as representing what is happening over an area. Specifically, in this case over all instruments the average IXH = 2.33 but with a minimum of 0.50 and a maximum of 3.54 over just this small area and over almost two hours.

These values can then be compared to those for the stratiform rain plotted in Fig. 10b. The entire rain field is WSS, so any spatial variations among the different detectors are not nearly as important as in the convective rain. For this rain the average IXH is only 0.002, while the minimum and maximum values are 0 and 0.018, respectively. Essentially, then, there is no variation of IXH in this stratiform rain.

c. *Evaluating* IXH *over larger areas*

So far the focus has largely been on time series observations. There are occasions, however, when it might also be useful to evaluate IXH over an area at one time, such as one scanned by a radar or perhaps over a large network of rain gauges.

However, before pursuing this, we first consider some synthesized data both in 1D and 2D simulations for which the inputs are known; that is, we know ahead of time whether the data are statistically heterogeneous or WSS. For the two 1D examples, we have expectations based upon the results in the previous sections. Indeed, Fig. 11 shows that these expectations are met in one dimension, giving credence to the methodology. The application of the technique, however, is a little more complicated in two dimensions as discussed next.

In this first example, a random field of rainfall rates is correlated (Fig. 12) using the square root matrix method as described in Jameson and Heymsfield (2014) and Jameson (2015). Any apparent "blocking" of the data is due to the spatial resolution used as discussed in Jameson and Heymsfield (2013, 2014) and Jameson (2015). This data field has a global mean value of 100 mm h^{-1}, although this particular value does not matter; 10 mm h^{-1} could have been used just as easily. First, let us evaluate the α statistics. This is done by proceeding along a continuous path in the north–south direction beginning at the top and then reversing direction at the bottom by moving over one column and then going back toward the top. This minimizes artificial jumps in the counting process at the boundaries of the domains. For convenience we will refer to this as our north–south path. However, there are two dimensions, so we repeat the process going east and west as well to produce our east–west path for a better sampling of the data field; the total path is the combination of the two paths. An LM curve (a least squares error or a local regression spline spanning twice the observed decorrelation distance) is fit and the fluctuation curve is calculated as before. The α values are then computed. Here we note that other paths are possible, of course. What is important is that whatever path is chosen it should span the

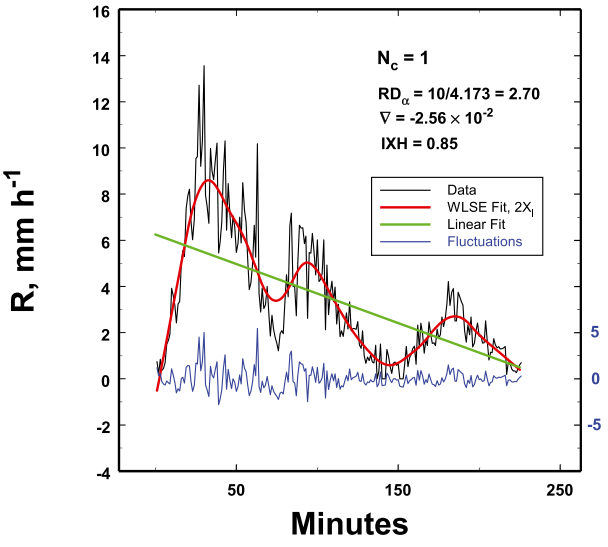

FIG. 9. A second set of data from Wallops Island is shown. The data are statistically heterogeneous because of the change in the variance of the fluctuations responsible for the significant α. The likelihood that this is a random occurrence is only about 2%.

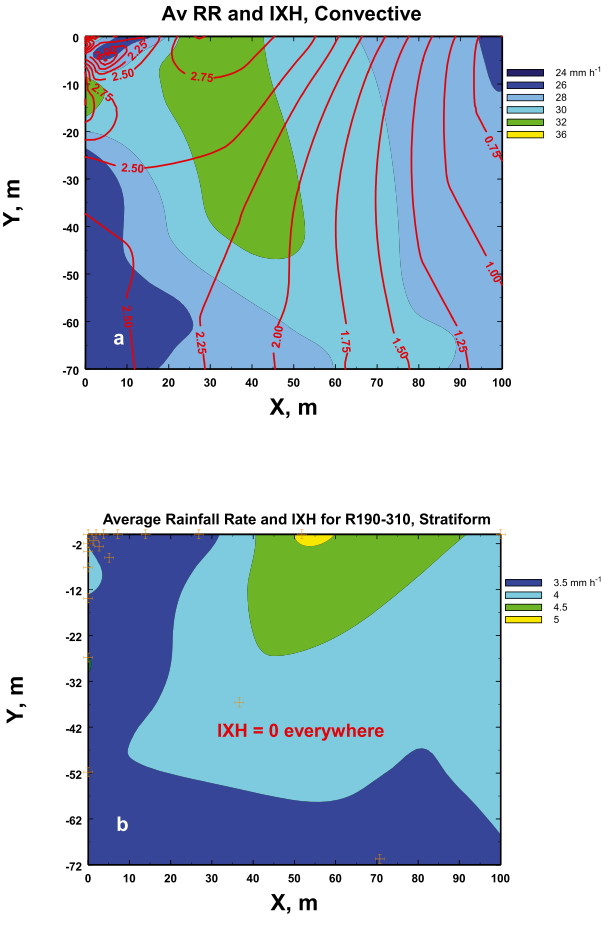

FIG. 10. (a) The 2D interpolated fields of the temporal average rainfall rate and IXH (red lines) for the convective rain during the 440-min rain event using each of the values of each instrument (orange crosses) separately. Clearly, depending upon location, there is significant variability in both the global mean rainfall rate and IXH, although all values of IXH are consistent with statistical heterogeneity. (b) The 2D interpolated fields of the temporal global mean rainfall rate and IXH for the stratiform rain during the 440-min rain event. The spatial variability is less than in (a), and the values of IXH over the entire area are consistent with statistical homogeneity.

entire domain, remembering that σ_α depends upon the total N. The statistical likelihood that the total path value of α is random can then be estimated using a normal distribution having zero mean as discussed previously. If one were dissatisfied with the result along one path, then one is always free to choose another as long as it spans the entire domain.

In Fig. 12a the total α over both paths is 0, while σ_α for 80 000 samples is 6.394. Consequently, $\mathrm{RD}_\alpha = 0$. Moreover, as expected from the data construction, the Bayesian inversion yields $N_c = 1$, so we conclude that the data are WSS (IXH = 0) in accordance with (5) and in agreement with the input conditions.

In Fig. 12b the same calculations are performed but for an input of statistically heterogeneous conditions. In this case $\mathrm{RD}_\alpha = 13/6.394 = 2.03$ while the Bayesian inversion yields $N_c = 2$, so that again the statistically heterogeneous data are correctly identified. Consequently, there is some justification for proceeding to the analyses of real 2D observations using this procedure.

Figure 13a is a plot of data collected in rain in Colorado on 12 September 2005 using the Colorado State University–University of Chicago–Illinois State Water Survey (CSU–CHILL) radar now operated in Greeley, Colorado, by Colorado State University for the National Science Foundation. The radar measured the radar reflectivity factor Z that has been converted into estimates of the rainfall rate using the relation of Sekhon and Srivastava (1971) [$R = (Z/300)^{0.741}$]. Applying the same approach as described above, $\alpha = 0$, so $\mathrm{RD}_\alpha = 0$. On the other hand, the Bayesian inversion reveals two significant peaks to the 99% significance level, that is, 99% of the distribution. Hence, IXH = 0.5, so these data are statistically heterogeneous, but they may be considered to be a statistical mixture because $\mathrm{RD}_\alpha = 0$.

As an another example, R is plotted in Fig. 13b for a storm on 6 July 2005 near Greeley, again observed using the CHILL radar and converting Z into R as described above. The total α was 13, so $\mathrm{RD}_\alpha = 13/5.675 = 2.291$. On the other hand, N_c was 3. The size of RD_α indicates that these data are statistically heterogeneous at the 99% level of statistical reliability even aside from the large N_c.

As a final example, we consider another case from 6 June 2005 as illustrated in Fig. 13c. For these data, $|\alpha| = 1$ with $\sigma_\alpha = 4.32$. Thus, $\mathrm{RD}_\alpha = 0.232$, while $N_c = 4$ (viz., 9, 18, 54,

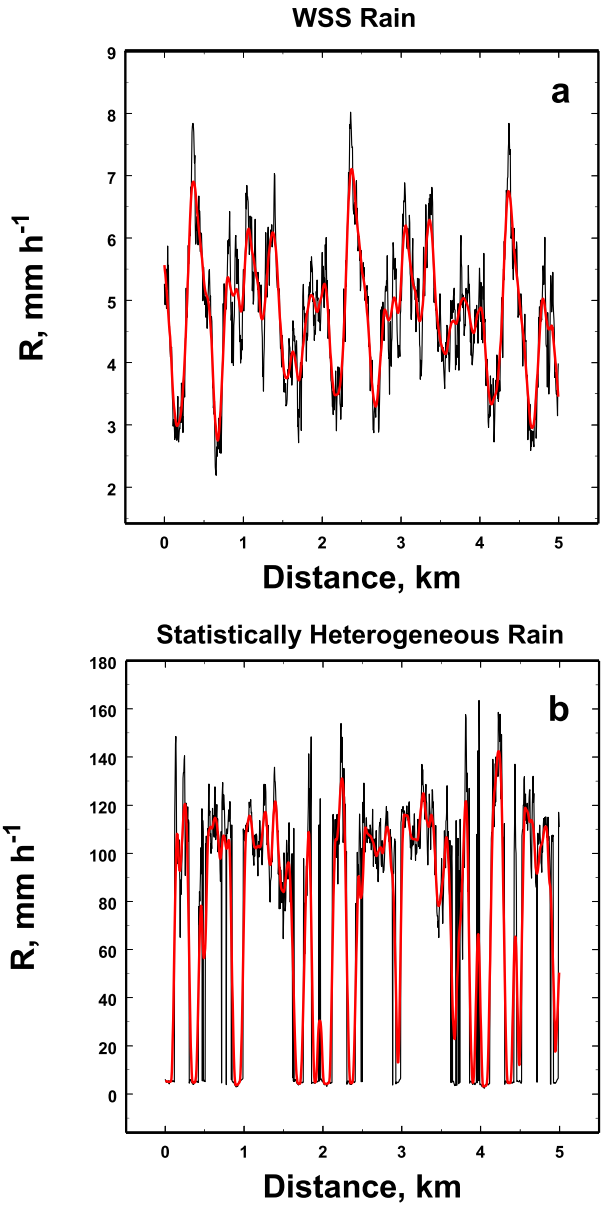

FIG. 11. Two one dimensional examples of synthesized data for (a) WSS data having $\alpha = 2$, $RD_\alpha = 0.41$, and $N_c = 1$ so that IXH = 0 and for (b) statistically heterogeneous data having $\alpha = 14$, $RD_\alpha = 2.90$, and $N_c = 2$, so that IXH = 1.70 in agreement with the input assumptions in both cases. As before, the red line denotes the WLSE local regression fit spanning twice the decorrelation length.

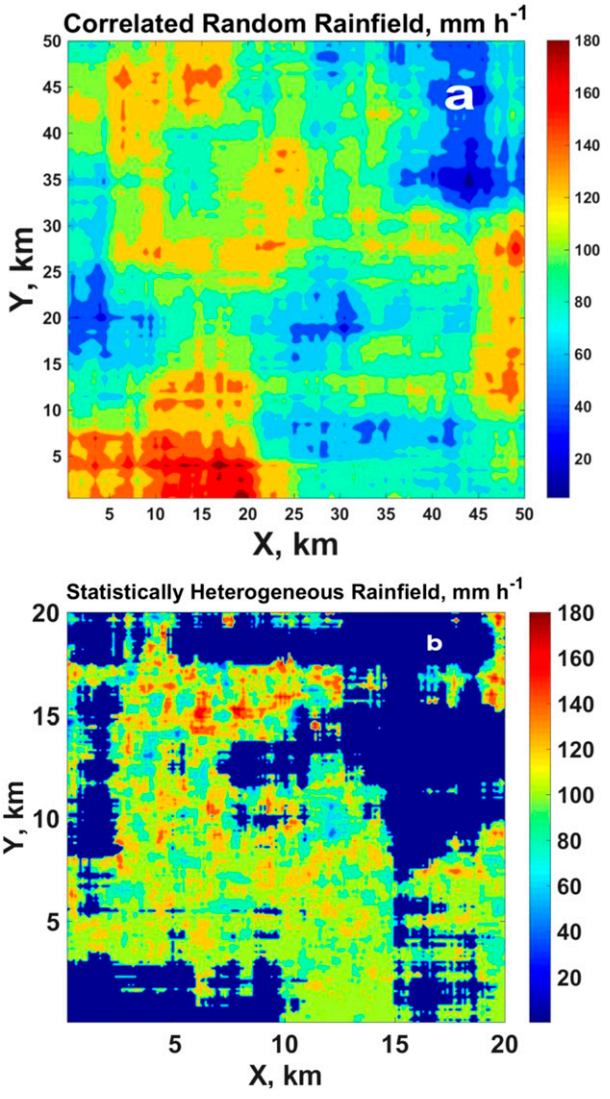

FIG. 12. Synthesized 2D rain fields (a) of a WSS rain field having $\alpha = 0$ and $N_c = 1$ so that IXH = 0 and (b) of a statistically heterogeneous rain field having $\alpha = 14$, $RD_\alpha = 2.03$, and $N_c = 2$ so that IXH = 0.77, which is in agreement with the assumed inputs in both cases.

and 108 mm h^{-1}). IXH is then 1.5 so that these data appear to be statistically heterogeneous with the heterogeneity being driven exclusively by the number of Bayesian components, that is, by the variability in the global mean values. However, the small value of the relative dispersion suggests that these data may also be a statistical mixture.

In summary, then, the validity of this approach is substantiated by numerical simulations of rain data such as in Figs. 11 and 12. More importantly, however, several samples in real rain of the application of this approach yielded seven examples of statistically heterogeneous rain and three examples of rain events that are WSS. This is still a small set of data, so the representativeness of these data remains to be determined through continued analyses. Statistically heterogeneous rain appeared to be mostly, but not exclusively, associated with convective rain. This statistical heterogeneity is sometimes driven by the variability in the global mean values and sometimes by the variability in the variance. Hence, both quantities must be monitored. This implies that the Wiener–Khintchine theorem may not always be applicable when attempting to derive scaling

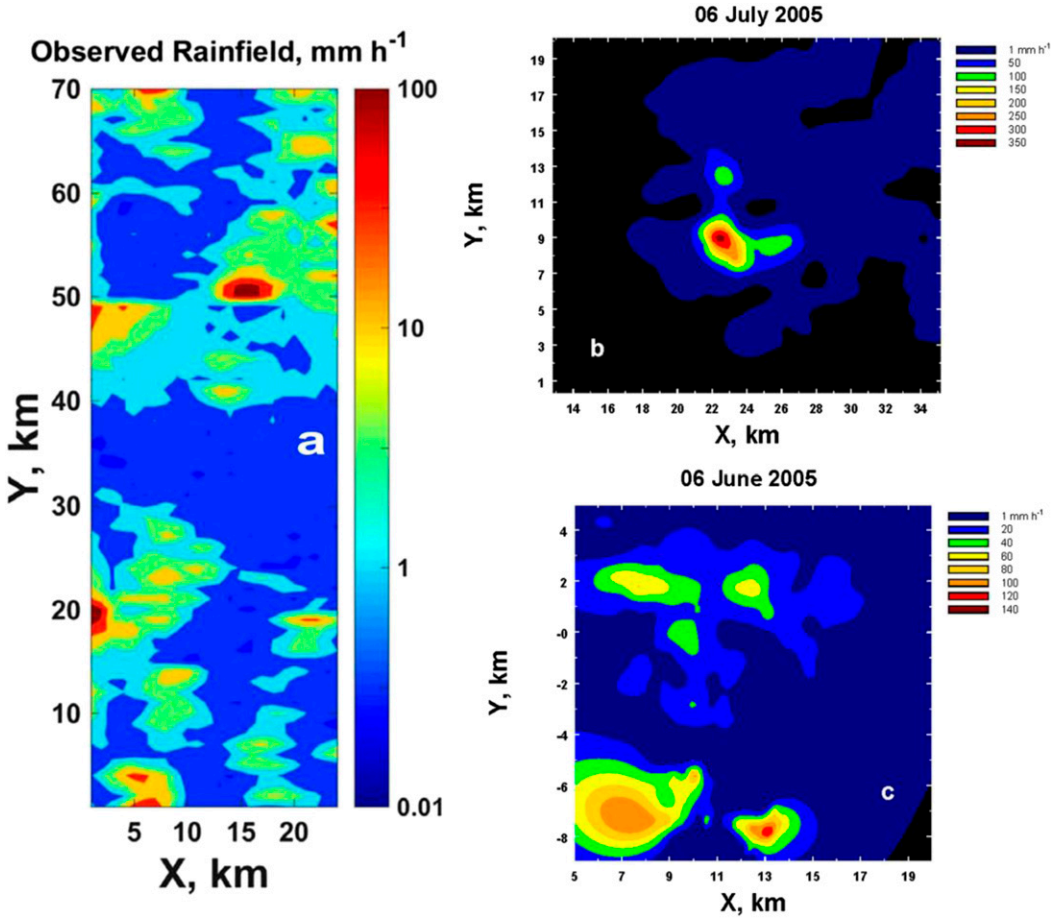

FIG. 13. Examples of rainfall rate estimated from the radar reflectivity. (a) For these data, $RD_\alpha = 0$ and $N_c = 2$. Hence, IXH = 0.5, so these data are statistically heterogeneous, but they may be considered to be a statistical mixture because $RD_\alpha = 0$. (b) In this case $RD_\alpha = 2.29$ and $N_c = 3$. Consequently, IXH = 1.26, so these data are statistically heterogeneous. (c) As in Fig. 12b, but for a different storm. While $RD_\alpha = 0.232$ is small, $N_c = 4$, so IXH = 1.5. These data, therefore, exhibit statistical heterogeneity because of different contributing components, but they may also almost be considered to be a statistical mixture.

relations for the rainfall rate. In statistically heterogeneous rain, the Fourier transform of the observed correlation functions need not correspond to the actual variance spectrum of the rain, so such derivations may be misleading, particularly for larger IXH.

4. Conclusions

Up to now there has been no model independent method for evaluating whether a set of data in rain was WSS. This work provides one such method for determining the validity of an assumption of WSS. An index of statistical heterogeneity (IXH) is developed such that as its value approaches zero, the more acceptable is the assumption of WSS and the more acceptable is the use of the Wiener–Khintchine theorem. This approach also has potential implications with regard to other commonly used techniques, such as kriging and autoregressive modeling, for example.

Several examples from real rain measurements demonstrate the potential applicability of this approach to actual observations of rain. It seems likely, too, that a statistically heterogeneous set of data can sometimes be subdivided into separate statistically homogeneous and statistically heterogeneous portions (e.g., see Figs. 3–6) so that the Wiener–Khintchine theorem can be applied to at least some of the data. It is also shown that when there is statistical heterogeneity, the value of IXH for a single instrument can depend on the location of that instrument even when a location may vary by only tens to hundreds of meters (Fig. 10).

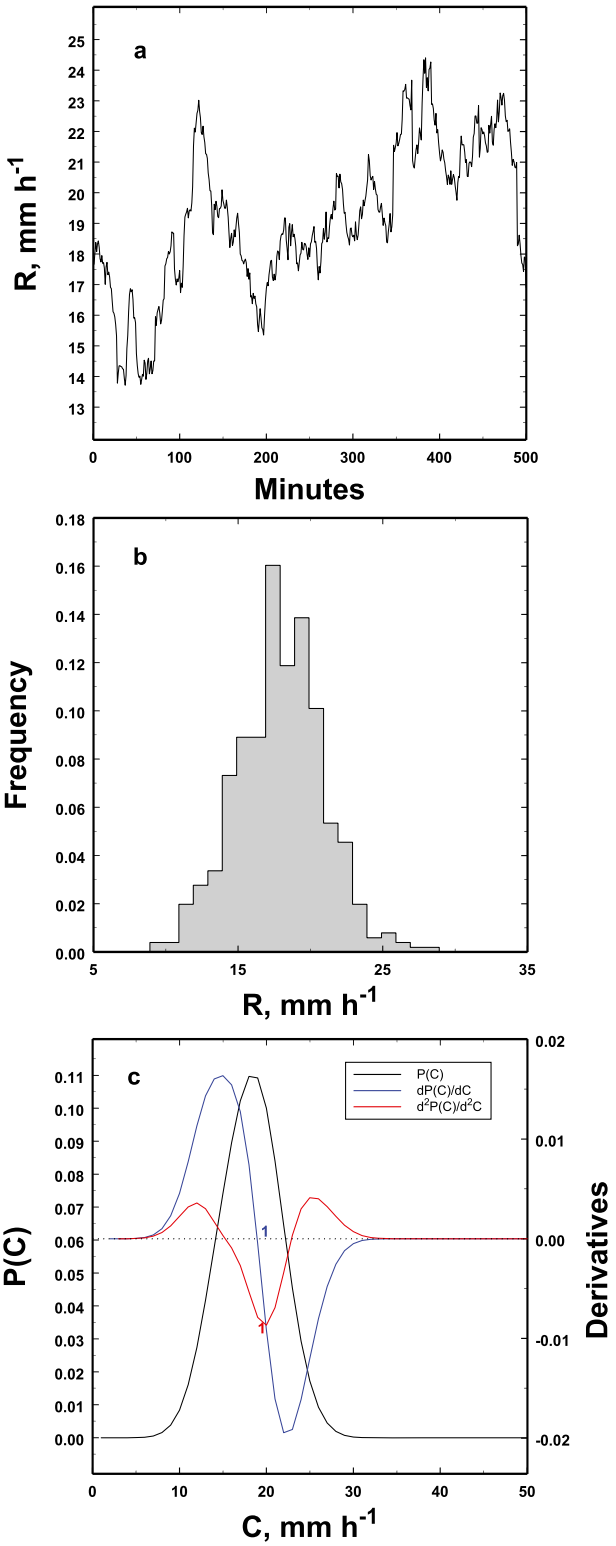

FIG. A1. (a) The time series plot of statistically homogeneous correlated rainfall rate shows no particular average value. (b) The frequency histogram for the data plotted in (a), suggesting one or two peak values. (c) The Bayesian inversion (black) distribution $P(C)$ of the global mean values C as well as the first (blue) and second (red) derivatives. The first derivative shows only one zero crossing, indicating one component as confirmed by the one minimum in the second derivative.

In addition to developing a broader base of examples in order to get a better feel for the frequencies and meteorological conditions associated with statistical homogeneity, heterogeneity, and mixtures, it is also necessary to study how the magnitude of IXH is related to the significance of any errors through the misapplication of the Wiener–Khintchine theorem—in other words, when does statistical heterogeneity become a serious problem and how serious is it? This is clearly a research project in itself that will be addressed in future work under preparation.

Acknowledgments. This work was supported (AJ) by the National Science Foundation (NSF) under Grants AGS1532423, AGS1532977 (ML), and AGS-1639868 (AK).

APPENDIX

Numerical Examples of Analysis Procedures

a. Resolution optimization for the Bayesian inversion procedure

Resolution is an important consideration in the Bayesian inversion process. When using the normal distribution for such inversions, the resolution is controlled by the standard deviation σ. The process of inversion is analogous to the binning of data for a histogram in that the optimum σ is one that involves a reasonable number of data points but not so many that it smooths out significant details of the distribution. For the binning of histograms, Freedman and Diaconis (1981) devised an algorithm that depends upon the width of the data distribution between the first and third quartile [interquartile range (IQR)] as well as the number of observations n, namely,

$$\sigma = \frac{2(\mathrm{IQR})}{\sqrt[3]{n}}. \tag{A1}$$

A multitude of blind numerical experiments (generated independently by one of the authors) over dozens of different scenarios having different numbers of contributing components, each having different fractional contributions to the mean values and different variance structures, revealed that for our purposes it is more appropriate with respect to the inversion process

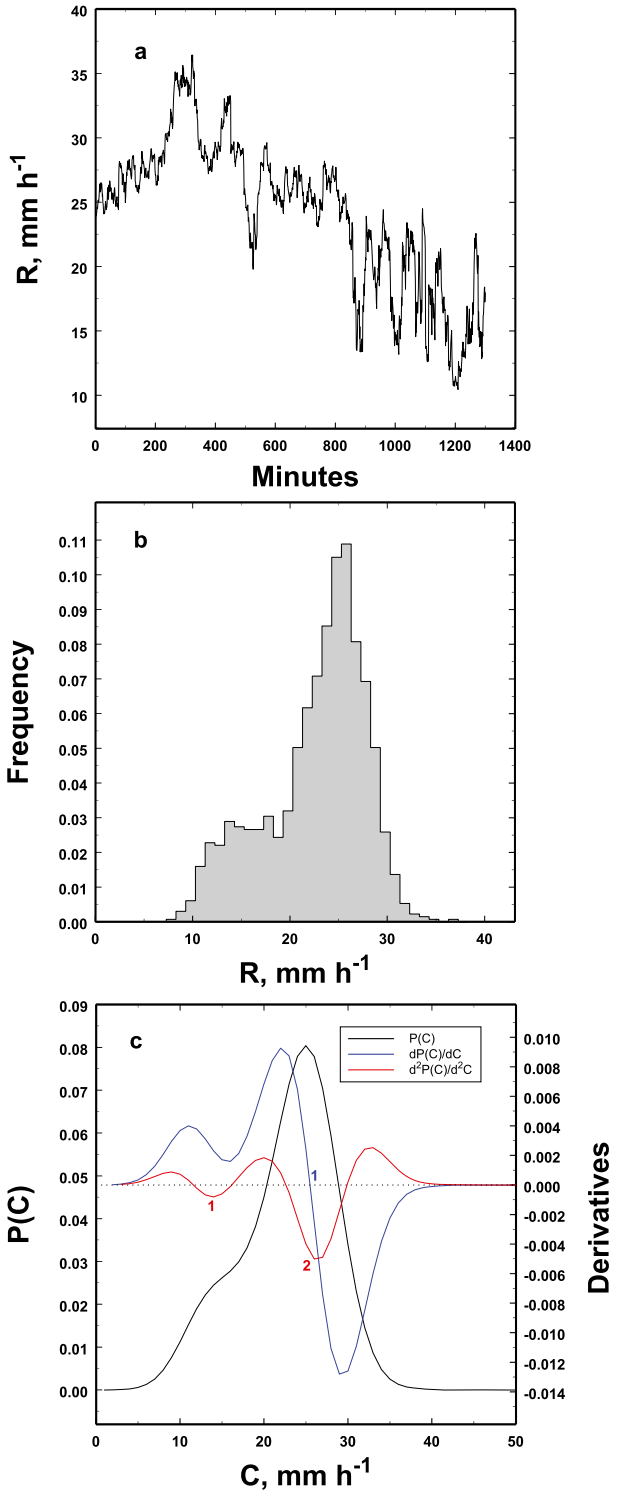

FIG. A2. As in Fig. A1, but for (a) a two-component statistically heterogeneous set of values. (b) The histogram clearly shows that there are likely two components. (c) However, the first derivative indicates only one component, but the second derivative detects the presence of both components, thus illustrating the importance of looking at the second derivate as well as the first derivative.

to double that value and then to require that $0.05 \leq \sigma \leq 5.0$. Values outside of either of these bounds were then set to the appropriate limits. Still in order to be detected successfully, it must be remembered that two or more components contributing to a distribution have to be sufficiently separated by an amount dependent upon their relative magnitudes. The approach used here appears to produce reasonable results over dozens of numerical tests and with respect to the observed data. It is also worth repeating that if components are so close so as to become inseparable, it is likely that any associated statistical heterogeneity will not preclude the successful application of the Wiener–Khintchine theorem; or, to put it another way, such data are effectively statistically homogeneous whether or not there are detectable individual components.

b. Examples of the component detection procedure

There are several approaches for detecting peaks indicating components. The first and most obvious one is to look for descending zero crossings of the first derivative. Sometimes, though, the peaks are hidden, so one must resort to looking at the second derivatives (e.g., Arteaga-Falconi et al. 2015; Slodzinski et al. 2013). Examples of the usefulness of this approach are illustrated below. Finally, if there is still a question, one can use the residual technique, that is, looking at the residual curve after subtracting fits for the other components from $P(C)$ as was done, for example, in Jameson (2007). This is also illustrated below.

1) AN EXAMPLE OF STATISTICALLY HOMOGENEOUS DATA

We begin by taking draws from a single normal distribution of R to produce a correlated time series shown in Fig. A1a. From this plot alone there is no way to tell how many components might be contributing to the dataset. This is not clear even looking at the frequency histogram given in Fig. A1b. However, the Bayesian inversion clearly shows only one component as confirmed by the single zero crossover for the first derivative of $P(C)$ and the single minimum in the second derivative.

2) AN EXAMPLE OF STATISTICALLY HETEROGENEOUS DATA

The results are plotted in Fig. A2. This time draws from two different normal distributions were used. Again, the time series in Fig. A2a reveals nothing about the components contributing to the data. However, in this case the histogram in Fig. A2b strongly suggests the presence of at least two components. The results of the Bayesian inversion (Fig. A2c) confirm this conclusion. While there is only one zero crossover for the first derivative, there are

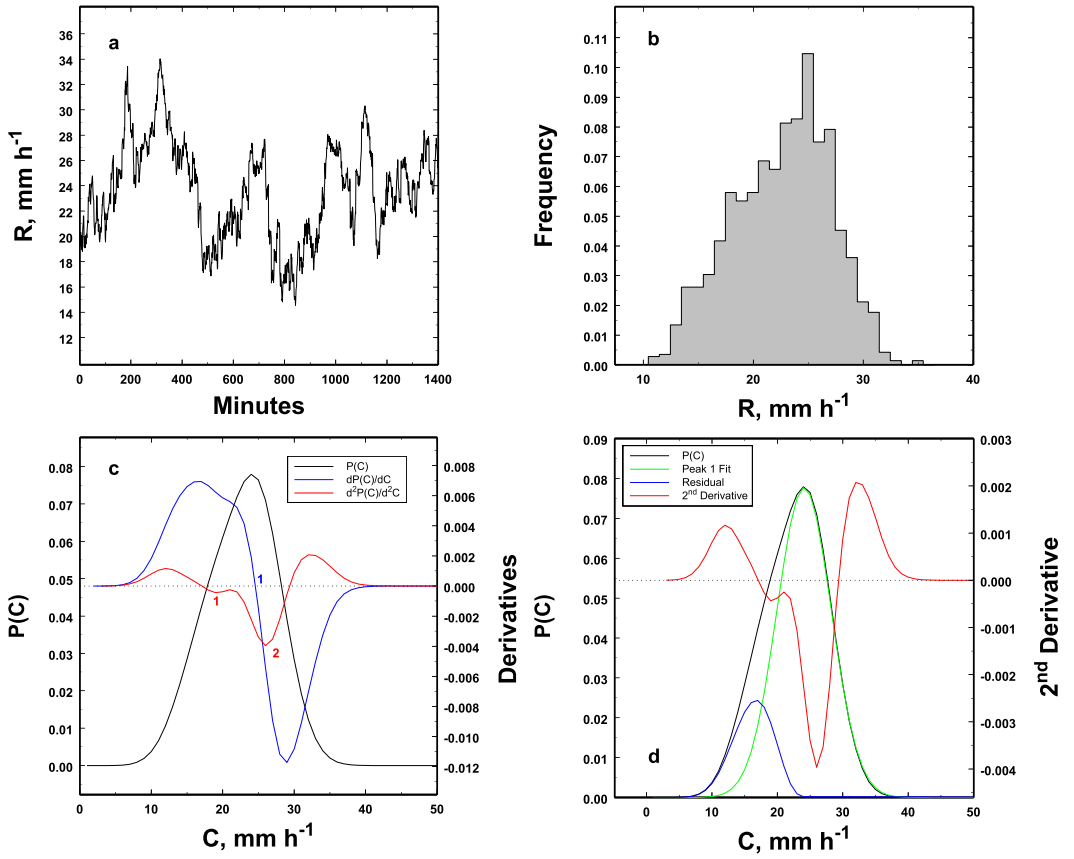

FIG. A3. (a) As in Fig. A2, but the two peaks are not so obvious in (b) the histogram. (c) However, while the first derivative detects only one component, the second derivative is still able to capture the presence of both components. (d) When the second derivative is indecisive, one can also resort to the technique of residuals, where the subtraction of one peak (green) from $P(C)$ leaves a residual peak (blue) that is consistent with the second derivative (red). At some point all techniques might fail, but then the data are likely effectively statistically homogeneous anyway.

two minima in the second derivative, indicating the presence of the two components, one of which is only sufficient to produce an inflection in $P(C)$. Often, however, the histogram is not all that helpful in guessing what might be happening as illustrated next.

3) A MORE SUBTLE EXAMPLE OF STATISTICALLY HETEROGENEOUS DATA

As in the previous example, draws from two different normal distributions are combined, leading to the results plotted in Fig. A3. This time not only is the time series (Fig. A3a) uninformative but neither is the histogram (Fig. A3b), which suggests a single or perhaps more components. Whether there is more than one component is not clear until we perform the Bayesian inversion (Fig. A3c). This time $P(C)$ almost looks as though it consists of a single component with a slight distortion on the left side of the distributions. The second derivative shows, however, that there is a second component. This plot also shows that attention must be paid to relative minor minima in the second derivative, especially when they appear to be associated with more unusual structures in the first derivative. It is also possible to use the residual detection technique of fitting the one component and subtracting that from $P(C)$ as illustrated in Fig. A3d. The residual clearly supports the detection by the second derivative of a second component.

REFERENCES

Anderson, A., and A. Kostinski, 2010: Reversible record breaking and variability: Temperature distributions across the globe. *J. Appl. Meteor. Climatol.*, **49**, 1681–1691, https://doi.org/10.1175/2010JAMC2407.1.

——, and ——, 2011: Evolution and distribution of record-breaking high and low monthly mean temperatures. *J. Appl. Meteor. Climatol.*, **50**, 1859–1871, https://doi.org/10.1175/JAMC-D-10-05025.1.

Arteaga-Falconi, J., H. Al Osman, and A. El Saddik, 2015: R-peak detection algorithm based on differentiation. *2015 IEEE Ninth*

International Symposium on Intelligent Signal Processing (WISP): Proceedings, IEEE, 30–33, https://doi.org/10.1109/WISP.2015.7139157.

Crane, R. K., 1990: Space-time structure of rain rate fields. *J. Geophys. Res.*, **95**, 2011–2020, https://doi.org/10.1029/JD095iD03p02011.

Freedman, D., and P. Diaconis, 1981: On the histogram as a density estimator: L_2 theory. *Z. Wahrscheinlichkeit. Verw. Geb.*, **57**, 453–476, https://doi.org/10.1007/BF01025868.

Glick, N., 1978: Breaking records and breaking boards. *Amer. Math. Mon.*, **85**, 2–26, https://doi.org/10.1080/00029890.1978.11994501.

Jameson, A. R., 2007: A new characterization of rain and clouds: Results from a statistical inversion of count data. *J. Atmos. Sci.*, **64**, 2012–2028, https://doi.org/10.1175/JAS3950.1.

——, 2015: A Bayesian method for upsizing single disdrometer drop size counts for rain physics studies and areal applications. *IEEE Trans. Geosci. Remote Sens.*, **53**, 335–343, https://doi.org/10.1109/TGRS.2014.2322092.

——, and A. J. Heymsfield, 2013: A Bayesian approach to upscaling and downscaling of aircraft measurements of ice particle counts and size distributions. *J. Appl. Meteor. Climatol.*, **52**, 2075–2088, https://doi.org/10.1175/JAMC-D-12-0301.1.

——, and ——, 2014: Bayesian upscaling of aircraft ice measurements to two-dimensional domains for large-scale applications. *Meteor. Atmos. Phys.*, **123**, 93–103, https://doi.org/10.1007/s00703-013-0303-3.

——, M. L. Larsen, and A. B. Kostinski, 2015: Disdrometer network observations of finescale spatial–temporal clustering in rain. *J. Atmos. Sci.*, **72**, 1648–1666, https://doi.org/10.1175/JAS-D-14-0136.1.

Jerison, D., I. M. Singer, D. W. Stroock, and N. Wiener, Eds., 1997: *Proceedings of Symposia in Pure Mathematics: The Legacy of Norbert Wiener: A Centennial Symposium*, Vol. 60, American Mathematical Society, 405 pp.

Khintchine, A., 1934: Korrelationstheorie der stationären stochastischen Prozesse. *Math. Ann.*, **109**, 604–615, https://doi.org/10.1007/BF01449156.

Kiely, G., and K. Ivanova, 1999: Multifractal analysis of hourly precipitation. *Phys. Chem. Earth*, **24B**, 781–786, https://doi.org/10.1016/S1464-1909(99)00080-5.

Nason, G. P., 2006: Stationary and non-stationary time series. *Statistics in Volcanology*, H. M. Madder et al., Eds., Special Publications of IAVCEI, No. 1, Geological Society, 129–142.

Schleiss, M., S. Chamoun, and A. Berne, 2014: Nonstationarity in intermittent rainfall: The "dry drift." *J. Hydrometeor.*, **15**, 1189–1204, https://doi.org/10.1175/JHM-D-13-095.1.

Sekhon, R. S., and R. C. Srivastava, 1971: Doppler radar observations of drop-size distributions in a thunderstorm. *J. Atmos. Sci.*, **28**, 983–994, https://doi.org/10.1175/1520-0469(1971)028<0983:DROODS>2.0.CO;2.

Serinaldi, F., C. Kilsby, and F. Lombardo, 2018: Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Adv. Water Resour.*, **111**, 132–155, https://doi.org/10.1016/j.advwatres.2017.10.015.

Sibson, R., 1981: A brief description of natural neighbor interpolation. *Interpreting Multivariate Data*, V. Barnett, Ed., John Wiley & Sons, 21–36.

Slodzinski, R., L. Hildebrand, and W. Vautz, 2013: Peak detection algorithm based on second derivative properties for two dimensional ion mobility spectrometry signals. *Integration of Practice-Oriented Knowledge Technology: Trends and Prospectives*, M. Fathi, Ed., Springer, 341–354, https://doi.org/10.1007/978-3-642-34471-8_28.

Watson, D. F., 1992: *Contouring: A Guide to the Analysis and Display of Spatial Data: With Programs on Diskette.* 1st ed. Pergamon Press, 321 pp.

Wiener, N., 1930: Generalized harmonic analysis. *Acta Math.*, **55**, 117–258, https://doi.org/10.1007/BF02546511.