The Effective Number of Spatial Degrees of Freedom of a Time-Varying Field

CHRISTOPHER S. BRETHERTON, MARTIN WIDMANN, VALENTIN P. DYMNIKOV,*
JOHN M. WALLACE, AND ILEANA BLADÉ

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 11 March 1998, in final form 14 August 1998)

ABSTRACT

The authors systematically investigate two easily computed measures of the effective number of spatial degrees of freedom (ESDOF), or number of independently varying spatial patterns, of a time-varying field of data. The first measure is based on matching the mean and variance of the time series of the spatially integrated squared anomaly of the field to a chi-squared distribution. The second measure, which is equivalent to the first for a long time sample of normally distributed field values, is based on the partitioning of variance between the EOFs. Although these measures were proposed almost 30 years ago, this paper aims to provide a comprehensive discussion of them that may help promote their more widespread use.

The authors summarize the theoretical basis of the two measures and considerations when estimating them with a limited time sample or from nonnormally distributed data. It is shown that standard statistical significance tests for the difference or correlation between two realizations of a field (e.g., a forecast and an observation) are approximately valid if the number of degrees of freedom is chosen using an appropriate combination of the two ESDOF measures. Also described is a method involving ESDOF for deciding whether two time-varying fields are significantly correlated to each other.

A discussion of the parallels between ESDOF and the effective sample size of an autocorrelated time series is given, and the authors review how an appropriate measure of effective sample size can be computed for assessing the significance of correlations between two time series.

1. Introduction

The number of effective spatial degrees of freedom (ESDOF) is among the most fundamental properties of a time-varying spatial field such as 500-mb height. It can serve as a basis for (i) estimating the number of stations or grid points required to represent the field (Bagrov 1969; TerMegreditchian 1969, 1990), (ii) assessing field significance (e.g., Lorenz 1969; Livezey and Chen 1983; van den Dool and Chervin 1986; Wallace et al. 1991), (iii) characterizing and comparing observed and simulated fields (Fraedrich et al. 1995), and (iv) estimating the dimension of an underlying attractor that may describe the evolution of the field (Toth 1995; Fraedrich et al. 1998, manuscript submitted to J. Atmos. Sci.; Dymnikov and Gritsun 1997). It could also be useful for assessing the feasibility and robustness of matrix inversions such as those used in statistical prediction.

---

* Current affiliation: Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow, Russia.

---

Corresponding author address: Dr. C. S. Bretherton, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640.
E-mail: breth@atmos.washington.edu

© 1999 American Meteorological Society
ments (for the variance functional, these are determined by the spatially integrated variance and kurtosis of the field itself); see, for example, Bagrov (1969) and TerMege6itchian (1969); visual inspection (Wallace et al. 1991); and the interquartile range (Toth 1995; Bagrov 1969). For a given functional, these approaches will give almost identical answers. An attractive feature of this measure of ESDOF is that if the fluctuations in the field variable are normally distributed, the ESDOF can be easily estimated from the eigenvalue spectrum of the temporal covariance matrix (TerMege6itchian 1969; Fraedrich et al. 1995).

Other measures of ESDOF have been developed for other problems. Livezey and Chen (1983) discussed how to test whether the correlation between two spatial fields, such as 500-mb height and sea surface temperature, is statistically significant. Their approach is to determine the fraction of the grid points at which the local temporal correlation is significant at some level, then to test whether this fraction is significantly larger than one would expect if the fields were independent. The critical fraction for field significance depends considerably on the spatial correlation structure of the fields and can be higher than if all grid points were independent. Livezey and Chen defined an ESDOF for their problem using a Monte Carlo approach, as did van den Dool (1987) for a similar type of problem.

The concept of ESDOF is certainly no less clearly defined (in a statistical context) and of no less practical value than the number of significant modes in an EOF expansion, yet it seems to be much less widely understood and appreciated in the research community. Its usage thus far has tended to be restricted to those who have, in effect, reinvented it and customized it for their own purposes.

This paper is intended for researchers who wish to use the ESDOF methodology without reinventing it. In section 2, we review Bagrov’s and TerMege6itchian’s formulas for ESDOF. The remaining sections go beyond prior work to systematically investigate the performance of these ESDOF formulas in contexts relevant to geophysical data analysis. Section 3 shows that for a typical application the ESDOF formulas give results similar to those of past investigators with much less computation. It also explores how long a time record is needed to accurately determine ESDOF, and how ESDOF can be affected by nonnormality of the data. In section 4 we demonstrate that these measures of ESDOF are appropriate for tests of field significance, for example, whether a realization of a field deviates significantly from its mean, or whether two realizations of a field are significantly correlated, even when the data are nonnormally distributed. Section 5 shows the connection of ESDOF to the parallel concept of effective sample size for temporal autocorrelated time series and briefly reviews the application of effective sample size to statistical significance testing of the correlation coefficient between two temporally autocorrelated time series. Section 6 presents our conclusions.

2. Effective number of degrees of freedom of a field

Following Bagrov (1969) and TerMege6itchian (1969), let us consider a time-varying spatial field defined at N stations and over some finite time interval, within which the mean value is identically equal to zero at all stations, and let $\psi_i(t)$ be its value at the $i$th station at time $t$. We choose some quadratic functional of the $\psi_i(t)$, such as the variance summed over all $N$ stations,

$$E(t) := [\psi(t), \psi(t)] := \sum_{i=1}^{N} \psi_i^2(t),$$

where $\psi(t)$ denotes the vector of the $\psi_i(t)$ and $(,)$ denotes an inner product. We use this to construct a measure of the effective number of spatial degrees of freedom. This measure, $N^*$, is defined as the number of uncorrelated random normal variables $a_i$, each having zero mean and the same population variance $(a^2)$, for which the $\chi^2$ distribution for the specified functional most closely matches the PDF of the functional of $\psi(t)$. This approach has been used by Lorenz (1969) and Wallace et al. (1991) as a basis for estimating the statistical significance of analogs. In order to increase the number of samples for estimating the PDF they used as their functional squared Euclidean distance, which the $\chi^2$ distribution match the PDF of the functional of $\psi(t)$.

Rather than determining the best match visually, Bagrov (1969) required that the $\chi^2$ distribution match the observed distribution’s ensemble mean value ($E$) and its temporal variance of $E$ about the ensemble mean,

$$\text{var}(E) = \langle E^2 \rangle - \langle E \rangle^2$$

(i.e., the first two moments). Unless otherwise noted, ensemble means ($\langle \cdot \rangle$) are estimated on the basis of time averages of all available data. For the $\chi^2$ distribution, $\langle E \rangle = N^*(a^2)$ and $\text{var}(E) = 2N^2(a^2)^2$. Solving for $N^*$, we obtain a “moment matching” estimate of ESDOF,

$$N^*_{\text{mm}} = \frac{2\langle E^2 \rangle}{\text{var}(E)}, \quad (a^2)_{\text{mm}} = \frac{\text{var}(E)}{2\langle E \rangle}.$$  \(1\)

Using this relationship it is possible to estimate $N^*$ directly from the time series of $E$ without explicit calculation of its PDF. There is no guarantee that $N^*_{\text{mm}}$ remains within the anticipated range $1 \leq N^*_{\text{mm}} \leq N$ unless the number of independent sampling times is large and the $\psi_i$’s are normally distributed.

An alternative way of estimating $N^*$, also discussed

---

1 We will use the term “station” to denote the locations in three-dimensional space at which the field is defined.
by Bagrov (1969) and TerMegreditchian (1969), is based on the \( N \times N \) covariance matrix \( \mathbf{C} \) of the \( \psi_i \)'s. This approach is based on two assumptions: (i) that the \( \psi_i \)'s are normally distributed and (ii) that \( \mathbf{C} \) is known with sufficient accuracy (i.e., that it is based on a relatively large number of independent samples). We assume that \( \mathbf{C} \) has eigenvalues \( \lambda_i \) and normalized principal components \( z_i(t) \) formed by projecting the field vector \( \psi(t) \) onto the \( k \)th normalized eigenvector (EOF) \( \mathbf{e}_k \).

We may now calculate \( N^* \) in terms of the \( \lambda_i \) from (1):

\[
E(t) = \sum_{k=1}^{N} \lambda_i z_i^2, \\
\langle E \rangle = \sum_{k=1}^{N} \lambda_k.
\]

and

\[
\text{var}(E) = \sum_{k=1}^{N} \lambda_i^2 \text{var}(z_i^2) = \sum_{k=1}^{N} \lambda_i^2 (\langle z_i^2 \rangle - \langle z_i \rangle^2)^2 \\
= \sum_{k=1}^{N} \lambda_i^2 (\langle z_i^2 \rangle - \langle z_i \rangle^2). 
\]

Since the \( z_i \) are normalized and normally distributed, it follows that \( \langle z_i^2 \rangle = 1, \langle z_i \rangle = 3 \), and

\[
\text{var}(E) = 2 \sum_{k=1}^{N} \lambda_i^2. 
\]

Note that for a time record of finite length or nonnormally distributed observations \( \psi_i(t) \), this estimate of \( \text{var}(E) \) will be different from that obtained directly from the time series of \( E \). Matching the mean and variance of a \( \chi^2 \) distribution with \( N^* \) degrees of freedom from these estimates as in (1), we obtain an “eigenvalue formula” estimate of ESDOF,

\[
N^*_\text{eff} = \left( \sum_{k=1}^{N} \lambda_i \right)^2 / \sum_{k=1}^{N} \lambda_i. 
\]

In fact, \( N^*_\text{eff} \) depends only on the partitioning of the variance between the EOFs. If \( f_k \) is the fraction of the overall variance explained by EOF \( k \), (3) can be written

\[
N^*_\text{eff} = \left( \sum_{k=1}^{N} f_i z_i^2 \right)^{-1}.
\]

It is readily verified that \( 1 \leq N^*_\text{eff} \leq N \), so that in contrast to \( N^*_\text{max} \), \( N^*_\text{eff} \) does remain within the anticipated range. In the limiting case of equal \( \lambda_i \), \( N^*_\text{eff} = N \), whereas if a single EOF accounts for all the variance of \( \psi_i \), \( N^*_\text{eff} = 1 \).

The \( N^*_\text{eff} \) can be estimated directly from \( \mathbf{C} \) without performing eigenvector analysis by rewriting (3) in the form

\[
N^*_\text{est} = \frac{\text{tr}(\mathbf{C})^2}{\text{tr}(\mathbf{C}^2)} = \frac{\sum_{i=1}^{N} C_{ii}}{\sum_{i,j=1}^{N} C_{ij}^2}.
\]

Here, \( \text{tr}(\mathbf{C}^2) \) is also the square of the Frobenius norm of \( \mathbf{C} \).

The estimates \( N^*_\text{est} \) and \( N^*_\text{eff} \) need not be identical unless the number of independent observations of the spatial field is large and the principal components are normally distributed.

In order to obtain some insight into the interpretation of \( N^*_\text{eff} \), suppose that the eigenvalues of the covariance matrix drop off geometrically with \( k \), that is,

\[
\lambda_{k+1}/\lambda_k = \exp(-\mu), \quad (\mu > 0).
\]

If \( \mu \) is large, just a few EOFs dominate the variance, while if \( \mu \) is small, the variance will be more evenly distributed among the EOFs. This type of variance partitioning frequently occurs in real geophysical applications (see Fig. 1 of section 3 for an example). Estimating \( N^*_\text{eff} \) based on (3) yields

\[
N^*_\text{est} = \frac{\sum_{k=0}^{\infty} \exp(-k\mu)}{\sum_{k=0}^{\infty} \exp(-2k\mu)} = \frac{1 + \exp(-\mu)}{1 - \exp(-\mu)}.
\]

As we would expect, \( N^*_\text{est} \) is large if \( \mu \) is small and many modes contribute significantly to the variance. When \( \mu \) is fairly small, \( N^*_\text{est} \approx 2/\mu \) and the fraction of the overall variance explained by the first \( N^*_\text{est} \) EOFs is

\[
\frac{\sum_{k=0}^{N^*_\text{est}-1} \exp(-k\mu)}{\sum_{k=0}^{\infty} \exp(-k\mu)} = 1 - \exp(-N^*_\text{est} \mu) \\
\approx 1 - e^{-2} = 86\%,
\]

regardless of the rate at which the eigenvalues decrease. Huang et al. (1996) mentioned that in their experience, roughly 90% of the variance was explained by the first \( N^*_\text{est} \) EOFs. They were computing ESDOF using a different method in a different context and did not present an argument why this should be so, but the agreement is reassuring.

In geophysical applications, the observations are often not normally distributed, and \( N^* \) must be estimated based on a limited number of sampling times. In the next two subsections, we discuss these complications, which affect \( N^*_\text{est} \) and \( N^*_\text{max} \) in different ways. We consider both \( N^*_\text{est} \) and \( N^*_\text{max} \), since we will see later that each can be useful for different purposes.
a. Differences between \( N_{ef}^* \) and \( N_{mm}^* \) due to non-normally distributed data

The moment-matching estimate \( N_{mm}^* \) of ESDOF is not based on any assumptions about the data at individual stations, although it should be expected that when the data are far from normally distributed, the actual distribution of \( E(t) \) should not be well approximated by a \( \chi^2 \) distribution, even if the two distributions have the same mean and variance. However, the assumption of normality did enter the derivation of the formula for \( N_{ef}^* \) in the calculation of \( \text{var}(E) \), where it was assumed that the fourth moment of the normalized principal components, also known as their kurtosis, is equal to 3. For observed distributions, the kurtosis can vary considerably. For a distribution with long tails, the kurtosis \( \kappa \) can be much larger than 3. For example, the exponential distribution has \( \kappa = 9 \). For a compact distribution, \( \kappa \) can be small, for example, \( \kappa = 1.8 \) for a uniform distribution.

To gain some sense how nonnormality affects \( N_{ef}^*/N_{mm}^* \), we assume that all the principal components (PCs) have the same kurtosis. Repeating the derivation of (3), we obtain

\[
\frac{N_{ef}^*}{N_{mm}^*} = \frac{\kappa - 1}{2}. \tag{9}
\]

This ratio can vary considerably; it is 4 for PCs with an exponential distribution and 0.4 for PCs with a uniform distribution.

b. Effect of sample length on estimated \( N_{ef}^* \) and \( N_{mm}^* \)

Estimates of both \( N_{mm}^* \) and \( N_{ef}^* \) based on a limited set of sampling times may differ considerably from their “true” (long sample record) values. To quantify these errors, we let a hat above a quantity denote an estimate of that quantity based on observations at a random set of \( T \) independent sampling times. Such an estimate is a random variable; we denote its mean by angle brackets and its standard deviation by \( \delta \) placed in front of it. We will use \( \langle N_{mm}^* \rangle \) and \( \delta N_{mm}^* \) to characterize the bias and scatter in \( N_{mm}^* \) compared to the true value \( N_{mm}^* \) and similarly for \( N_{ef}^* \).

We are not aware of any prior analysis of the bias and scatter of \( N_{mm}^* \). Fraedrich et al. (1995) estimated the scatter of \( N_{ef}^* \) by using North et al.’s (1982) rule of thumb for the typical error in eigenvalues of a covariance matrix determined from a limited set to observations:

\[
\delta N_{ef}^*/N_{ef}^* = 4(2/T)^{1/2}. \tag{10}
\]

Fraedrich et al. (1998, manuscript submitted to J. Atmos. Sci.) also found empirically that \( N_{ef}^* \) tend to be smaller than the true \( N_{ef}^* \) but did not propose a formula for this low bias. One can understand how temporal undersampling can lead to a low bias of \( N_{ef}^* \) by considering the limiting case of a single sample of a zero mean process, \( T = 1 \). If the one measured value at station \( i \) is \( \phi_i \), the elements of the covariance matrix are just \( \hat{C}_{ij} = \phi_i \phi_j \). The covariance matrix can be written

\[
\hat{C} = \phi \phi^T,
\]

where \( \phi \) is the vector with components \( \phi_i \). From the spectral theorem (Strang 1988, p. 296), such a matrix has rank 1, that is, one nonzero eigenvalue, with corresponding eigenvector \( \phi \). Hence, the estimated \( N_{ef}^* \) is 1, which is always an underestimate of \( N_{ef}^* \). Similarly, an estimate \( \hat{N}_{ef}^* \) based on \( T \) observations can never exceed \( T \).

The same principle can be expected to hold for other estimates of ESDOF. To distinguish \( N^* \) independent spatial modes, we must be able to differentiate between their temporal variability. We cannot expect to do this accurately unless there are many more observing times than modes, that is, unless \( T \gg N^* \).

In the appendix, we derive formulas for the bias and scatter of estimates of \( N_{ef}^* \) and \( N_{mm}^* \) in the parameter regime

\[
T \gg N^* \gg 1. \tag{11}
\]

This excludes samples so short (\( T \leq N^* \)) that one cannot expect a useful estimate of \( N^* \) and also excludes fields with a small number (in practice, 5 or less) of ESDOF. If \( N^* \) is small, our formulas are still qualitatively useful, but may be in error by a factor of 2 or more.

For \( \hat{N}_{ef}^* \) we find

\[
\langle \hat{N}_{mm}^* \rangle /N_{mm}^* \approx 1 + 2/T. \tag{12}
\]

\[
\delta \hat{N}_{mm}^*/N_{mm}^* \approx (2/T)^{1/2}, \tag{13}
\]

For short samples, \( \hat{N}_{mm}^* \) has a slight positive bias, but the error is dominated by the scatter.

For \( \hat{N}_{ef}^* \) we find

\[
\langle \hat{N}_{ef}^* \rangle /N_{ef}^* \approx \frac{1}{1 + N_{ef}^*}, \tag{14}
\]

\[
\delta \hat{N}_{ef}^*/N_{ef}^* \approx c T^{-1/2}, \tag{15}
\]

Unlike Fraedrich’s formula (10), which was obtained using a scaling analysis, these formulas are rigorously derived. The constant \( c \), given in (A26), depends upon the eigenvalue spectrum. It is typically less than 0.5, much less than predicted by Fraedrich’s formula, so the scatter of \( \hat{N}_{ef}^* \) is very small. In fact, if all EOFs explained the same fraction of the variance, or if one EOF explained the entire variance, \( c \) would be zero. The \( N_{ef}^* \) has a considerable low bias unless \( T \) is many times as large as \( N_{ef}^* \). The predicted bias is 50% for \( T \approx N_{ef}^* \) independent samples and 11% for \( T \approx 10 N_{ef}^* \) independent samples.

Let us consider a specific example. Consider a field observed at a large number of stations with a normal distribution of observed values at each station, and \( N^* = 20 \) ESDOF. With an infinitely long time record, we would correctly conclude that \( N_{mm}^* = N_{ef}^* = 20 \). With
by normalizing the time series at each station to have the same (e.g., unit) variance, which may be regarded as an implicit form of weighting. This can slightly improve the power of statistical tests of field significance.

In this case, (3) assumes the form

$$N_{ef} = \frac{N^2}{\sum_{i} r_{ij}^2}$$

where $r_{ij}$ is the correlation coefficient between $\psi_i$ and $\psi_j$. This formula was also derived by Fraedrich et al. (1995).

If the covariance matrix $C$ is accurately known, one can define a “prewhitening” norm,

$$E(t) = [\psi(t), C^{-1}\psi(t)] = \sum_{k=1}^{N} z_k^2(t).$$

The second equality shows that this is equivalent to defining $E(t)$ using the standard metric, but using the normalized principal components $z_k(t)$ in place of the individual station data. The $z_k(t)$ are uncorrelated and all have the same variance of 1. Hence one obtains $N_{ef}$ = $N$ degrees of freedom, the maximum possible. One might regard this norm as attractive for maximally discriminating between two realizations of a process. However, it should be used with great caution, as it weights PCs that have very small variance (and that are probably subject to large measurement uncertainty) as highly as the dominant PCs.

3. Application to real datasets

In this section, we discuss the estimation of ESDOF for two meteorological datasets. Our focus is on the following issues.

1) How similar are $N_{ew}^w$ and $N_{ef}^w$ in typical applications?
2) How robust are estimates of $N_{ew}^w$ and $N_{ef}^w$ from a limited set of times and locations?
3) Can we identify from the data situations in which we should be particularly cautious in interpreting and using ESDOF?

a. Northern Hemisphere wintertime 500-mb heights

Our primary example is a dataset of Northern Hemisphere wintertime 500-mb heights averaged over 5-day intervals for 48 winter seasons (DJF of 1946–93), based on National Meteorological Center (now the National Centers for Environmental Prediction) daily analyses. The data are given on a 5° × 7.5° lat–long grid northward of 20°N, at a total of 672 grid points. Gridpoint values are area-weighted when calculating spatial averages, and at all grid points, data are converted into anomalies by subtracting the mean annual cycle at that grid point. This dataset is the same as was used by Wallace et al. (1991) but extended for 9 yr.

Figure 1 shows the fraction of the variance explained

![Figure 1. Cumulative fraction of the variance explained by the leading EOFs of the 500-mb height pentads.](image-url)
by the leading EOFs. Many EOFs contribute significantly to the variance; the leading 10 modes explain only about 68% of the overall variance and 27 modes are required to explain 90%. The frequency distribution of the two leading PCs are shown in Fig. 2. They exhibit noticeable positive skewness and have kurtosis of 2.6 and 2.7, respectively, both slightly less than the kurtosis of the normal distribution. This is characteristic of all 10 leading PCs. Hence, based on (9), we should expect that the moment matching estimate $N^{*}$ of ESDOF will be slightly larger than the eigenvalue formula $N_{eff}^{*}$.

Figure 3 shows the two estimates of ESDOF $N^{*}$ and $N_{eff}^{*}$, computed based on data from samples of different lengths. Recall that we use a hat to distinguish an estimate made on the basis of a finite set of sampling times. In the top panel, the samples consist of consecutive intervals. In the bottom panel, they are nonconsecutive and uniformly spaced throughout the entire data record to better ensure independence of the samples. For each possible period of a given sampling length, for example, one season = 18 pentads, the ensemble means and standard deviations of the estimates from all possible samples (each of the 48 winters, in this case) are shown. The theoretical estimates from section 2b of the mean and standard deviation of the estimates are also shown for comparison.

We see that in this dataset, for large sample sizes, $N^{*}_{max} = 24$ is indeed systematically larger than $N_{eff}^{*} = 19$, due to the slight nonnormality of the leading PCs. We have also performed the same calculation as shown in Fig. 3 with the actual time series of the normalized PCs replaced by independent samples drawn from a standard normal distribution. This does not change the distribution of variance between the EOFs, so it does not change $N^{*}_{max}$. However, it reduces $N^{*}_{max}$ so that it is indistinguishable from $N_{eff}^{*}$ for large sample sizes as expected. The small sample bias and scatter of $N^{*}_{max}$ and $N_{eff}^{*}$ are unchanged. Figure 4 shows $\chi^2$ distributions fit to the observed frequency distribution of spatially averaged variance $E(t)$ with $N^{*}_{max} = 24$ and $N_{eff}^{*} = 19$ degrees of freedom. The distribution based on the moment-matching estimate gives a better fit, reflecting the bias in $N_{eff}^{*}$ deriving from the assumption of normality. However, Wallace et al. (1991), who computed the frequency distribution of the covariance and correlation coefficient between all pairs of maps for the same dataset (over a slightly smaller number of years), found that fitted distributions with 20 ($=N^{*}_{eff}$) degrees of freedom (DOF) best matched the observed distributions. This result also held for our longer data record. In the next section, we
Fig. 4. $\chi^2$ distributions with $N_{\text{min}}^*=23.8$ (solid) and $N_{\text{y}}^*=18.5$ (dashed) degrees of freedom superposed on the observed frequency distribution of spatially averaged 500-mb height variance.

will see that in general, $N_{\text{min}}^*$ ESOF gives a good fit to variance distributions, while $N_{\text{y}}^*$ gives a good fit to covariance and correlation distributions; that is, the two measures of ESOF are useful in different contexts.

It is also interesting to note that the first $N_{\text{y}}^*$ EOFs explain about 82% of the variance, which is quite close to the prediction (8) of 86% that was derived assuming that the variance fraction associated with the EOFs decreases geometrically with EOF index. For the dominant EOFs, the assumption of geometric falloff is quite good for this example, though this approximation underestimates the variance fraction in the long tail of small-variance EOFs.

For the distributed samples, the theoretical formulas of section 2b for the bias and scatter of $\hat{N}_{\text{min}}^*$ and $\hat{N}_{\text{y}}^*$ agree very well with the bias and scatter calculated directly from the data. For moderate sampling lengths, $\hat{N}_{\text{min}}^*$ exhibits considerable scatter, increasing with decreasing sample size $T$ proportional to $T^{-1/2}$. When consecutive samples are used, $\hat{N}_{\text{min}}^*$ also tends to have a high bias for short samples. Since this effect nearly disappears when the samples are distributed through the entire time record, we attribute the high bias to temporal correlation of the 500-mb height fields between pentads. For a short sample, temporal correlation (e.g., a tendency toward anomalously low or high 500-mb heights over the North Pacific may persist over periods of weeks, months, or even years) causes the sample variance of $\hat{E}(t)$ to underestimate the true variance, which from (1) leads to an overestimate of $N_{\text{min}}^*$. There are also smaller, but noticeable, changes in the bias of $\hat{N}_{\text{y}}^*$ when consecutive samples are used.

Fraedrich et al. (1995) considered the effect of spatial resolution on $N_{\text{y}}^*$. They found that as long as the spacing between grid points is 500 km or less, $N_{\text{y}}^*$ computed for daily 500-mb height data is insensitive to resolution since the leading EOFs vary only on much larger length scales. We repeated this type of analysis for our pentad data (which is spatially smoother) and found that $N_{\text{y}}^*$ and $N_{\text{min}}^*$ are insensitive to spatial resolution even with a grid spacing as large as 1500 km.

b. A precipitation dataset

As our second example, we consider a dataset of daily precipitation at 113 rain gauge locations throughout Switzerland, for the winter seasons DJF 1961–90 (Widmann and Schär 1997). In this dataset, most (80%) of the variance is concentrated in the three leading EOFs. The frequency distribution of the PCs for the leading two EOFs is shown in Fig. 5. PC1 is highly positively skewed with a long tail and far from normal, as would be the rainfall distribution at any individual rain gauge. PC2 is not so skewed, but it also has very long tails.

Fraedrich et al. (1995) considered the effect of spatial resolution on $N_{\text{y}}^*$. They found that as long as the spacing between grid points is 500 km or less, $N_{\text{y}}^*$ computed for daily 500-mb height data is insensitive to resolution since the leading EOFs vary only on much larger length scales. We repeated this type of analysis for our pentad data (which is spatially smoother) and found that $N_{\text{y}}^*$ and $N_{\text{min}}^*$ are insensitive to spatial resolution even with a grid spacing as large as 1500 km.

Fig. 5. Frequency distributions of the leading two PCs of daily Swiss wintertime precipitation.
$N_{mm}^*$, DOF does satisfactorily fit the upper tail of the variance distribution. Since $N_{mm}^*$ is less than one, it is obviously not literally interpretable as a number of degrees of freedom. The value of $N_{mm}^*$ is a better indicator of the number of dominant spatial modes in the system.

If monthly rainfall totals are analyzed instead, $N_{mm}^*$ rises to 0.9, while $N_{ef}^*$ drops to 1.7. The monthly average rainfall frequency distribution is quite skewed, but less so than the daily data. As longer averaging intervals are used, we would expect the frequency distribution of rainfall to tend toward a normal distribution and the two estimates of ESDOF to converge.

4. Use of ESDOF for significance testing

An attractive application of ESDOF is to assess whether two realizations of a field are either significantly similar or significantly different in a statistical sense. For example, we may wish to compare the field with some prediction of that field at that time. Alternatively, we might want to compare the spatial structure of the field at two different times or a particular realization of a field with its mean value.

Wallace et al. (1991) [and earlier, Lorenz (1969), van den Dool and Chervin (1986), and others] have discussed ways of measuring the similarity or difference between two realizations of a field. We might compute the rms difference between the realizations, averaged over all grid points. Alternatively, we might compute the spatial correlation between the realizations. We expect that if the field has only a few ESDOF, then two randomly chosen realizations of the field will have a much broader distribution of spatial correlations and rms differences than if the field has many ESDOF. In particular, we might hope based on the examples shown in the previous section that the statistics of the field could be approximately modeled with a combination of $N^*$ independent random variables with standard normal distributions. This would permit classic statistical tests based on the chi-squared or Student's t-distributions to be applied to these problems. To justify this approach, we must determine whether the resulting PDF provides a conservative estimate of the frequency of outliers; this will be the goal of this section.

An assumption of most of this section is that the dominant principal components of the EOF decomposition of the field $\psi$ are approximately normally distributed, so that both the moment-matching and eigenvalue formulas for ESDOF are equal to a common value $N_{mm}^*$. This will be true if the gridpoint data are normally distributed, but not, for example, for the Swiss rainfall dataset. At the end of the section, we consider examples with highly nonnormal principal components and find that our statistical tests derived for normal PCs are still useful, but that the appropriate estimate of ESDOF is a combination of $N_{mm}^*$ and $N_{ef}^*$ that depends on the test.

a. Difference of a realization from the mean

As discussed in section 2, the PDF of the spatially integrated variance of a field, $E(t)$, is the generalized chi-squared distribution (2), that is, a variance-weighted sum of squares of unit normal random variables. In that section, we matched the mean and variance of this distribution to those of a standard chi-squared distribution with $N^*$ degrees of freedom, scaled by a variance factor $\langle a^2 \rangle$, hoping that the PDFs of the resulting distributions would also approximately match:

$$E(t)/\langle E \rangle \sim \chi^2(N^*)/N^*. \tag{16}$$

Here, the notation $\sim$ means “having approximately the same PDF as.” If these PDFs match reasonably well, we can use the 95th percentile of $\chi^2(N^*)$ to derive a confidence interval for $E(t)$.

We now investigate how well the PDFs of $E(t)$ and the $\chi^2$ distribution match in three diverse examples.

1) MANY EOFs WITH COMPARABLE VARIANCE

We might expect the PDFs to match quite well if many of the PCs explain comparable fractions of the overall variance. An idealized example is shown in Fig. 7.
POSE THAT AN EOF ANALYSIS OF A TIME SERIES OF A FIELD OBSERVED AT 10 GRID POINTS YIELDS 10 PRINCIPAL COMPONENTS WITH VARIANCES $\lambda_k = 0.19, 0.17, \ldots, 0.01$ (SINCE THESE VARIANCES SUM TO 1, THEY ARE ALSO THE VARIANCE FRACTIONS $f_k$ ASSOCIATED WITH THE 10 MODES). NO SUBSET OF THE PCs IS DOMINANT. IF EACH PRINCIPAL COMPONENT IS NORMALLY DISTRIBUTED, $E(t)$ HAS A GENERALIZED $\chi^2$ DISTRIBUTION GIVEN BY (2). FROM (3), WE FIND THAT THERE ARE $N^* = 7.5$ ESDOF.


THE EXACT (GENERALIZED $\chi^2$) AND MATCHED $\chi^2(N^*)$ DISTRIBUTIONS HAVE VIRTUALLY IDENTICAL CDFS. CLOSE EXAMINATION SHOWS THAT THE $\chi^2(N^*)$ DISTRIBUTION HAS A VERY SLIGHTLY LOWER PROBABILITY OF EXTREMELY LARGE VALUES, AND A VERY SLIGHTLY HIGHER PROBABILITY OF EXTREMELY SMALL VALUES THAN THE EXACT DISTRIBUTION, BUT THESE DIFFERENCES ARE SO SMALL THAT A STATISTICAL TEST CAN BE CONFIDENTLY BASED ON THE MATCHED $\chi^2$ DISTRIBUTION WITH $N^*$ ESDOF.

2) A FEW DOMINANT EOFs

AN OPPOSITE EXTREME IS IF ONE EOF IS DOMINANT, SUCH AS IN THE EXAMPLE SHOWN IN FIG. 8. IN THIS EXAMPLE, WE CONSIDER A FIELD MEASURED AT 11 POINTS FOR WHICH ONE EOF EXPLAINS HALF THE VARIANCE, AND THE OTHER 10 EOFs EACH EXPLAIN ONLY 5% OF THE VARIANCE. THIS IS A PARTICULARLY DIFFICULT CASE FOR MATCHING TO AN EFFECTIVE $\chi^2$ DISTRIBUTION, SINCE THE TAILS OF THE EXACT DISTRIBUTION WILL BE HEAVILY DOMINATED BY CONTRIBUTIONS FROM THE ONE LARGE EOF, WHILE THE EFFECTIVE CHI-SQUARED DISTRIBUTION MUST WEIGHT ALL ITS CONTRIBUTING NORMAL RANDOM VARIABLES EQUALLY.

THE MATCH IS LESS GOOD, BUT STILL REASONABLE, ESPECIALLY IN THE CRUCIAL UPPER TAIL. THE EXACT AND $\chi^2(N^*)$ DISTRIBUTIONS HAVE ALMOST IDENTICAL 95TH PERCENTILES AND AGREE CLOSELY EVEN FURTHER OUT INTO THE UPPER TAIL. THE DISCREPANCIES IN THE LOWER TAIL, THOUGH LARGER, ARE NOT THAT IMPORTANT FOR SIGNIFICANCE TESTING, SINCE WE WOULD BE UNLIKELY TO WANT TO TEST WHETHER A GIVEN REALIZATION WAS SIGNIFICANTLY CLOSER THAN EXPECTED TO THE MEAN.

WE ALSO EXAMINED A CASE WITH A FEW DOMINANT EOFS, AS COMMONLY FOUND IN CLIMATE STATISTICS APPLICATIONS. WE TOOK $\lambda_k = 0.5, 0.25, \ldots, 2^{-i}, \ldots.$, TRUNCATING AFTER 10 EOFs. THE RESULTS (NOT SHOWN) ARE SIMILAR TO THE CASE OF ONE DOMINANT EOF, BUT WITH CONSIDERABLY CLOSER AGREEMENT BETWEEN THE EXACT AND $\chi^2(N^*)$ DISTRIBUTIONS.

OUR CONCLUSION IS THAT THE MATCHED $\chi^2(N^*)$ DISTRIBUTION IS A REMARKABLY ACCURATE AND CONVENIENT TOOL FOR TESTING THE SIGNIFICANCE OF DEPARTURES OF A FIELD FROM ITS MEAN, AS LONG AS THE LEADING PRINCIPAL COMPONENTS OF THE FIELD ARE APPROXIMATELY NORMALLY DISTRIBUTED.

b. Difference of two realizations

ARE TWO REALIZATIONS OF A FIELD SIGNIFICANTLY MORE OR LESS DIFFERENT THAN WOULD BE TWO RANDOMLY CHOSEN RE-
alizations, that is, are they “analogs” or “antilogs”? One measure of this is the spatially integrated, squared difference between the two realizations, normalized by dividing by \( \langle E \rangle \). If their normalized principal components are \( y_k \) and \( z_k \), with difference \( \delta_k \), this difference can be written

\[
\Delta = \sum_{k=1}^{N} f_k \delta_k^2, \tag{17}
\]

where, as before, \( f_k \) is the variance fraction explained by EOF \( k \). To test this for statistical significance, we can compare \( \Delta \) to the percentiles of the distribution of the difference between two independent, randomly chosen instances of the field, for which \( \delta_k \) would have mean zero and variance equal to 2 (the sum of the variances of \( y_k \) and \( z_k \)). Thus \( 2^{-n} \Delta \), like the normalized spatially integrated variance \( E \), is distributed like a generalized chi-squared random variable with parameters \( f_k \), and its PDF is similar to that of \( \chi^2(N^*)/N^* \).

The results of the previous subsection imply that a significantly larger than expected \( \Delta \) can be tested for using a one-sided chi-squared test with \( N^* \) DOF. To test for a smaller than expected \( \Delta \), one can also use a one-sided test based, for instance, on the 5th percentile of a chi-squared distribution with \( N^* \) DOF. Since the 5th percentile of the exact distribution is slightly farther from zero than for the equivalent chi-squared distribution, a test based on \( \chi^2(N^*) \) gives a conservative significance threshold for \( \Delta \).

c. Field covariance and correlation

Another measure of the similarity of two realizations of a field is their “anomaly correlation,” that is, the spatial correlation of their departures from the time mean of the field. Unfortunately, we have not found a simple way to calculate the PDF of this correlation. We can, however, argue that this PDF should be quite similar to that based on a \( t \) test with \( N^* \) DOF, so that again an approximate significance test can be based on a standard test statistic.

Our argument is based on a calculation of the PDF of the covariance of two independent, randomly chosen realizations of the field. We normalize the covariance by the variance of the field. With notation as before, this covariance is

\[
C = \sum_{k=1}^{N} \left( f_k/2 \right) (r_k^2 - s_k^2). \tag{18}
\]

We can write \( y_k z_k \) as follows:

\[
y_k z_k = \left[ (y_k + z_k)^2 - (y_k - z_k)^2 \right]/4. \tag{19}
\]

If \( y_k \) and \( z_k \) are independent standard normal random variables, their normalized sum \( r_k = 2^{-1/2} (y_k + z_k) \) and difference \( s_k = 2^{-1/2} (y_k - z_k) \) are also independent standard normal random variables. Hence, we can write \( C \) as a generalized chi-squared random variable,
The correlation coefficient is the spatial covariance normalized by dividing by the square root of the product of the variances of the two realizations. Note that unlike in time series analysis, the covariance is computed without subtracting the spatial mean anomaly from each realization. This is because the null hypothesis is that both realizations are random cases of a field whose mean value at each grid point, and hence whose true spatial mean, is accurately known from our long time record of realizations of that field. Consider the PDF of the $R$ between two randomly chosen realizations of a field. Since both the variances and the covariance are well represented using distributions based on $N^*$ ESDOF, the same should hold for $R$. Hence, we anticipate that the PDF of $R$ has an approximate $t$ distribution with $N^* - 1$ DOF (we use $N^* - 1$, rather than $N^* - 2$, because the time mean of the field is known, adding one DOF to the classic test in which the means of each of the two correlated fields must be estimated as a sample mean).

Wallace et al. (1991) showed that for the 500-mb height dataset, a PDF of correlation generated from a $t$ distribution based on $20 (=N^*-2)$ ESDOF matched their observed PDF of $R$ very closely. One would deduce from this that two realizations of a 500-mb height field (or a forecast and an analysis of 500-mb height) are positively correlated at a 95% (one sided) significance level if $R > 0.37$.

Here we examine the two examples presented earlier in this section. We use 10 000 trials of Monte Carlo simulation of the principal component amplitudes to generate the “exact” distribution of $R$. For the example with many dominant modes (Fig. 10), the $t$ distribution matches the Monte Carlo simulation nearly perfectly. For the example with one dominant mode, the $t$ distribution shows a considerably higher probability of large correlations than the Monte Carlo simulation (Fig. 11). The $t$ test would produce conservative estimates of significance levels. In no example that we looked at did the $t$ test based on $N^* - 1$ DOF overestimate the significance of a particular correlation, so it is safe, and in most cases quite accurate, to use this test as long as time series of the dominant PCs are approximately normally distributed.

d. Nonnormally distributed data

How well do these significance tests based on $N^*$ degrees of freedom work if the grid points, and hence the PCs, are not normally distributed? The two extreme examples discussed below show that even in this case, the tests can still give good significance bounds, as long as $N^*$ is chosen appropriately for the test, as summarized in Table 1. For all of the tests, the appropriate $N^*$ lies somewhere in the range between $N^*_{\text{min}}$ and $N^*_{\text{opt}}$.

First, we consider the test discussed in section 4a for significance of the difference of a realization from the mean. This test is based on the CDF of the spatially integrated variance $E(t)$. We examine this CDF in the two extreme cases that the PCs are (a) exponentially distributed and (b) uniformly distributed, rather than normally distributed. These are examples of distribu-

### Table 1. ESDOF appropriate for various statistical tests.

<table>
<thead>
<tr>
<th>Significance test</th>
<th>Appropriate $N^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference of realization from mean</td>
<td>$N^*_{\text{min}}$</td>
</tr>
<tr>
<td>Difference of two realizations</td>
<td>$(2/N^<em>_{\text{min}} + 2/N^</em>_{\text{opt}})^{-1}$</td>
</tr>
<tr>
<td>Covariance and correlation</td>
<td>$N^*_{\text{opt}}$</td>
</tr>
</tbody>
</table>
Fig. 12. As in Fig. 7, but with the individual principal components (a) exponentially distributed and (b) uniformly distributed.

We consider one of the examples discussed earlier in this section, that of 10 principal components with variances 0.19, 0.17, ..., 0.01. With normally distributed PCs, Fig. 7 showed that the exact CDF of $E(t)$ was quite well approximated by a $\chi^2$ distribution with $N^* = 7.5$ ESDOF. Figure 12 shows the corresponding results for exponentially and uniformly distributed PCs. This does not affect $N^*_y$, which depends only on the partitioning of the variance between the EOFs. However, the moment-matching estimate $N^*_{mm}$ is much lower for the exponentially distributed PCs and much higher for the uniformly distributed PCs. This is because, as discussed in section 2, $N^*_{mm}$ differs from the eigenvalue estimate of 7.5 by a factor related to the kurtosis. Hence, the $\chi^2$ distribution that is matched to the exact distribution of $E(t)$ is quite different in these two cases, with a much broader variation around its mean for the exponentially distributed case. For the exponentially distributed PCs, the fit of $\chi^2(N^*_{mm})$ is only fair for small values of variance but is very good in the upper tail. The 95th percentile is very well estimated. The overall fit is nearly as good for the uniformly distributed PCs as for normally distributed PCs throughout the distribution. In both cases, the $\chi^2$ distribution based on the eigenvalue formula is an unsatisfactory fit to the true distribution. Both of these cases are quite extreme compared to most geophysical datasets, but the fit based on $N^*_{mm}$ ESDOF is good enough so that an approximate significance test may safely be based on $\chi^2(N^*_{mm})$.

Second, we consider the test discussed in section 4b for significance of the normalized difference $D$ of two realizations. The PDF of $D$ is determined by the PDFs of the differences $d_k = y_k - z_k$ between the PCs in the two realizations. These differences have a PDF that is intermediate between the PDFs of one realization and a normal distribution. In fact, if $y_k$ and $z_k$ are independently chosen from a distribution with kurtosis $\kappa$, then the kurtosis $\kappa_d$ of $d_k$ can easily be shown to be $(3 + \kappa)/2$, which is halfway between the kurtosis of the individual distributions and the kurtosis of a normal distribution. Hence, applying the moment-matching arguments of section 2 to estimate the ESDOF of $D$, we find from (9) that

$$N^*_D = 4N^*_y/(1 + \kappa_d) = (2/N^*_y + 2/N^*_{mm})^{-1}$$

is the harmonic mean of $N^*_y$ and $N^*_{mm}$.

Last, we consider the normalized covariance $C$. Equation (21) gave the approximate distribution of $C$ in terms of the difference of two $\chi^2(N^*)$ random variables. We can use moment matching to calibrate this formula for non-normally distributed PCs. The mean of $C$ is zero, so we match the variance. From (18), noting that $y_k$ and $z_k$ are mutually independent and have unit variance, we find

$$\text{var}(C) = \text{var} \left( \sum_{k=1}^{N} f_k y_k z_k \right) = \sum_{k=1}^{N} f_k^2 \text{var}(y_k) \text{var}(z_k) = \sum_{k=1}^{N} f_k^2 = 1/N^*_y,$$

that is, that now the eigenvalue estimate of $N^*$ may be the appropriate one. We might anticipate that the same would hold for the distribution of the correlation coefficient. Indeed, Fig. 13 plots the CDF of the correlation coefficient between two independent realizations of a
fewer spatial degrees of freedom but did not suggest a
of the Monte Carlo distribution. They suggested that the
ificant gridpoint correlations exceeded the 95th percentile
between the fields "significant" if the fraction of signif-
correlations are taken over years. They called the overall correlation be-
ed from different randomly
other. A Monte Carlo method for determining this dis-
structure as the given fields but are not related to each
realizations of the second field from different randomly
Andrews. By moment-match-
given fields and the variance is partitioned among the 10 EOFs
in the time series. We... are observed at grid points $t_1, \ldots, T$ and $j = 1, \ldots, N$, respectively, at times $t_1, \ldots, T$. Assume further that their gridpoint values have been normalized to zero mean and unit standard deviation. Let
be the empirical matrix of correlation coefficients be-
tween the two fields at all possible pairs of grid points,
based on the $T$ observations. A measure of the field
correlation between $\phi$ and $\psi$ is their integrated squared
cross covariance (ISCC):
\[
\text{ISCC} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (\hat{\beta}_{ij}^2).
\]
The ISCC lies between 0 and 1. If a lot of cross-cor-
relation coefficients are large, then the ISCC will also
be large. If the ISCC is close to one, then the covariances
between $\phi$ and $\psi$ explain almost all the variance of the
individual fields. The ISCC plays a similar role to Liv-
ezey and Chen's fraction of significant correlation co-
efficients. However, it also includes correlation of
at each location with $\psi$ at other locations, that is, telecon-
nections between the two fields.

Consider two fields, $\phi$ and $\psi$, that are unrelated. For
any finite sampling period, their empirical cross cor-
relations will still be nonzero, so their ISCC will be
positive. If their ISCC is computed from $T$ independent
samples, one may easily show that its expected value
is
\[
\langle \text{ISCC} \rangle = T^{-1}.
\]
Routine but lengthy computation shows that for large $T$,
\[
\text{var(ISCC)} = 2/(N_{\phi \psi} T^2),
\]
where
\[
N_{\phi \phi}^* = 2N_{\phi \psi} N_{\psi \psi}^*,
\]
and $N_{\phi \phi}^*$ and $N_{\psi \psi}^*$ are the ESDOF for the two fields computed using the eigenvalue formula. By moment-matching arguments, the distribution of the ISCC can be approximated by $\chi^2(N_{\phi \phi}^* T)$, allowing one to compute approximate confidence intervals for testing whether the ISCC is significantly larger than one would expect by chance, that is, whether the two fields are significantly correlated. Since the DOF will typically be very large, the $\chi^2$ distribution will have a very narrow peak, and an ISCC only fractionally larger than $T^{-1}$ will be statistically significant. Hence, a key issue in practical
use of this significance test (as with Livezey and Chen’s) is deciding whether observations at successive times are really independent and, if not, how best to estimate the true effective number of independent observation times.

5. Relation to effective sample size for serial correlation

There is a close connection between formulas for the ESDOF of a field and notions of effective sample size (ESS) for temporally correlated time series. This section has two purposes. First, we will show the connection between ESS and ESDOF. Second, we will mention some formulas in the literature for ESS for testing significance of correlations between time series. This is a common problem in atmospheric science, yet the relevant formulas and issues are mainly buried in old articles in statistical journals and are not widely appreciated.

The connection between ESS and ESDOF can be seen by applying the arguments of section 2 to $T$ successive observations from a zero-mean unit variance stationary Gaussian random process $X_t$, rather than to a realization of a spatial field. Different “realizations” would now correspond to independent sets of $T$ observations. For two realizations to be independent, each set of observations would have to be taken at a sufficiently long time from other sets so that temporal correlations between observations from different sets would be negligible. In developing the theory of ESS below, the temporal correlation structure is initially treated as a known. However, in practice it must be estimated. When we considered the ESDOF of spatial fields, the spatial covariance matrix could be accurately estimated from a few hundred independent observations of the spatial field. In contrast, we have only the one time series from which to estimate its correlation structure and ESS. This would be impossible were it not for the stationarity of the time series, which drastically simplifies its correlation structure. The estimation uncertainty in ESS may be considerable, particularly if the time series has substantial spectral power at periods of $T$ and longer.

Consider the distribution of the sample variance of the $T$ observations,

$$\langle E \rangle = \sum_{k=1}^{T} \lambda_k,$$

$$\text{var}(E) = 2 \sum_{k=1}^{T} \lambda_k^2.$$

Moment-matching implies that $E\langle E \rangle$ has an approximate distribution $\chi^2(T^*)T^*$ where $T^*$ is defined by the eigenvalue formula (3) and can be regarded as the effective sample size of the $T$ observations after accounting for their temporal correlation.

The $\lambda_k$ are sometimes called the singular spectrum $\lambda_k$. In practice, they must be estimated. When we considered the ESDOF of spatial fields, the spatial covariance matrix could be accurately estimated from a few hundred independent observations of the spatial field. In contrast, we have only the one time series from which to estimate its correlation structure and ESS. This would be impossible were it not for the stationarity of the time series, which drastically simplifies its correlation structure. The estimation uncertainty in ESS may be considerable, particularly if the time series has substantial spectral power at periods of $T$ and longer.

Consider the distribution of the sample variance of the $T$ observations,

$$E := \sum_{t=1}^{T} X_t^2.$$

The value of $E$ will depend on which $T$ observations were taken, so it has a PDF, whose mean and variance can be computed as in section 2. Let $\lambda_k$ be the eigenvalues of the lag-correlation matrix, whose elements are the lag correlations

$$C_{ij} = \langle X_i X_j \rangle = \rho_{i-j},$$

Then

$$T^* = \frac{\sum_{t=1}^{T} C_{ii}^2}{\sum_{t=1}^{T-1} \left(1 - |r|/T \right) \rho_i^2},$$

This formula dates back in the statistical literature at least to Bayley and Hammersley (1946) and is mentioned in some current texts on spectral analysis, for example, Percival and Walden (1993, p. 254).

Following the approach of section 4b of looking at the distribution of the sample covariance between two unrelated time series, one can derive an analogous formula for the ESS appropriate for significance tests of the correlation between two time series $X_t$ and $Y_t$, with different autocorrelation sequences $\rho_X^2$ and $\rho_Y^2$.

$$T_{XY}^* = \frac{T}{\sum_{t=1}^{T-1} \left(1 - |r|/T \right) \rho_X^2 \rho_Y^2},$$

Orcutt and James [1948, Eq. (4)] quote an equivalent formula. Davis (1976, p. 252) rederived this formula in a climate statistics application. The special case of this formula for red-noise (Markov) processes was derived even earlier by Bartlett (1935). If $r_X$ is the lag-one autocorrelation of $X$, and similarly for $Y$, and if $r_X, r_Y \ll 1$, then
Formulas (29) and (30) are sometimes said to define an ESS of order 2 (Kikkawa and Ishida 1988), as they are appropriate for significance tests involving second-order moments of the original time series (variance, covariance, and correlation). Formulas for computing the variance of the sample mean (i.e., the first-order moment) of T observations from a time series can be expressed in terms of a different ESS (of order 1); see, for example, Jones (1975) and Thiébaux and Zwiers (1984). For instance, for red noise with lag-one autocorrelation r, the ESS of order 1 is \( T^* = T(1 - r)/(1 + r) \) [see Jones (1975), which gives the discrete time version of Leith’s (1973) commonly used ESS formula], while (31) shows that the ESS of order 2 is \( T^* = T(1 - r^2)/(1 + r^2) \), which is up to twice as large as the Jones–Leith ESS when r approaches 1, that is, if the autocorrelation is strong.

The above definitions of ESS involves the actual and usually unknown covariance structure of the random process. Now we discuss estimation of the ESS from an observed time series. An approximation of ESS to within 10%–20% will suffice for most practical purposes. Several approaches have been proposed, all requiring care on the part of the user. First, we can use the empirical autocorrelations to estimate the true autocorrelations in the ESS formulas. The “biased” estimator of autocorrelation (Percival and Walden 1993, p. 191) should be used. To reduce the noise in this estimator if necessary, one can multiply the resulting autocorrelation sequence by a lag window (Percival and Walden 1993, p. 191) should be used. To reduce the noise in this estimator if necessary, one can multiply the resulting autocorrelation sequence by a lag window (Percival and Walden 1993, p. 191)

\[
T_{xy}^* = \frac{T - r_1 r_2}{1 + r_1 r_2},
\]

6. Conclusions

We have comprehensively investigated two measures of the effective number of spatial degrees of freedom in a time-varying field of data, calculated respectively from moment-matching \( (N_{em}^*) \) and from an eigenvalue formula \( (N_{ef}^*) \) based on EOF analysis, specifically the partitioning of variance between the EOFs. The two measures are the same if the data at each station are normally distributed and the field has been sampled at a large number of independent (serially uncorrelated) times. Each has both strengths and weaknesses, depending on the application.

We find that for short samples (of length less than a few hundred independent sampling times) estimates of \( N_{em}^* \) exhibit large scatter. Estimates of \( N_{ef}^* \) tend to be biased low unless the number of samples is many times as large as \( N_{em}^* \). This bias is due to insufficient sampling to distinguish the temporal variability of the different spatial modes. In general, our experience is that if the number of independent observation times is much larger than either of these measures of \( N^* \) (in practice, 10\( N^* \) or more), then the sample statistics are sufficient to define all of the dominant spatial modes (EOFs) accurately, but if the number of observation times is comparable to \( N^* \), then some leading EOFs will not be accurately determined.

If the field exhibits substantial autocorrelation from one sampling time to the next, the estimated \( N_{em}^* \) is biased high for short samples. If the frequency distribution of gridpoint data has long tails, such as for precipitation data, \( N_{em}^* \) is considerably smaller than \( N_{ef}^* \), while if the tails are truncated compared to a normal distribution, \( N_{ef}^* \) is smaller than \( N_{em}^* \); we quantify this in terms of the kurtosis of the principal components. We find that both of these measures are useful for statistical significance testing. For testing whether a given realization of a field is significantly different from the mean, we find that a chi-squared test with \( N_{em}^* \) degrees of freedom works surprisingly satisfactorily regardless of how the variance is partitioned between the EOFs and the probability distribution of the station data. For testing the significance of the correlation between two realizations of the field, such as a forecast and a verification, we show that a t test with \( N_{ef}^* - 1 \) degrees of freedom works well. A method is also proposed for testing whether two fields with different spatial correlation structure have a significant simultaneous correlation to each other. This leads to a chi-squared test with DOF dependent on the product of the \( N_{ef}^* \)’s of the two fields.

Last, we have briefly discussed the close analogy between effective spatial degrees of freedom of a time-varying field and effective sample size (ESS) or temporal degrees of freedom of a time series, that is, treating time instead of space as the dimension over which correlations occur. Formulas for ESS appropriate for assessing the significance of the correlation coefficient between two time series already exist and are in fact equivalent to our eigenvalue formula applied to the temporal autocorrelation matrix. In fact, in real applications involving fields of data, temporal and spatial autocorrelation both occur simultaneously. As long as the temporal autocorrelation structures of all dominant spatial modes are comparable, we can separate the two issues,
computing an ESSDFO relevant to spatial correlation and an ESS for temporal correlation.

Acknowledgments. The first author was supported by the National Science Foundation (NSF) under Grant DMS-9524770, the second by a grant from the Swiss National Science Foundation, and the third by the NSF Climate Dynamics Program under Grant ATM 9215512. Participation of the fourth author was sponsored jointly by NOAA (Office of Atmospheric Research and Office of Global Programs) and the University of Washington through the Joint Institute for the Study of the Atmosphere and Ocean. We are indebted to the Swiss Meteorological Institute–MeteoSwiss for providing access to their precipitation data. This work was partly motivated by stimulating discussions with George Gruza, Hans von Storch, Ian Jolliffe, and an anonymous reviewer also provided valuable advice.

APPENDIX

Bias and Scatter of ESDOF Estimates Based on Short Records

In this appendix, we derive formulas for the bias and scatter of the estimators $N_{\text{mean}}^*$ and $N_{\sigma r}^*$, which are based on a finite number $T$ of independent observation times, in the regime $T \gg N^* \gg 1$.

a. Bias and scatter in $N_{\text{mean}}^*$

We start from the definition (1) of $N_{\text{mean}}^*$. In this definition, we estimate the time mean and the variance of $E$ based on $T$ independent observation times:

$$
\hat{E} = \frac{1}{T} \sum_{i=1}^{T} E(t_i),
$$

$$
\hat{\text{var}}(E) = \frac{1}{T-1} \sum_{i=1}^{T} [E(t_i) - \langle E \rangle]^2.
$$

(A1)

While $E(t_i)$ will not in general have an exactly normal distribution, the central limit theorem states that the mean of $T$ independent samples of $E$ (and of its squared deviation from the mean) will have a distribution that is much closer to normal, since $T \gg 1$. Thus we regard $\hat{E}$ and $\hat{\text{var}}(E)$ as being normally distributed and entirely characterized by their means (denoted by square brackets) and standard deviations (denoted by the prefix $\delta$). These are easily calculated:

$$
\langle \hat{E} \rangle = \langle E \rangle,
$$

(A2)

$$
\delta \hat{E} = (\text{var}(E)/T)^{1/2},
$$

(A3)

$$
\langle \hat{\text{var}}(E) \rangle = \text{var}(E).
$$

(A4)

$$
\delta [\hat{\text{var}}(E)] \approx \text{var}(E)(2/T)^{1/2}.
$$

(A5)

Since $T$ is large, we have ignored the small relative difference between $T - 1$ and $T$ in computing the standard deviation of the variance. The relative standard deviations of $\hat{E}$ and $\hat{\text{var}}(E)$ can be found from the above:

$$
a = \delta \hat{E} / \langle \hat{E} \rangle = \left[ \text{var}(E) / \langle E \rangle \right]^{1/2} = (N_{\text{mean}}^* T)^{-1/2},
$$

(A6)

$$
b = \delta [\hat{\text{var}}(E)] / \langle \hat{\text{var}}(E) \rangle = (2/T)^{1/2}.
$$

(A7)

Note that both relative standard deviations are $O(T^{-1/2})$, which is small since $T$ is large.

It is convenient to normalize $\hat{E}$ and $\hat{\text{var}}(E)$ into the unit normal random variables $u$ and $v$, in terms of which we can write:

$$
\hat{E} = \langle E \rangle (1 + au) / \text{var}(E) (1 + bv),
$$

$$
\hat{\text{var}}(E) = \text{var}(E) (1 + bv).
$$

Substituting these expressions into the definition (1) of $N_{\text{mean}}^*$, we obtain

$$
\hat{N}_{\text{mean}}^* = \left( \langle E \rangle^2 / \text{var}(E) \right) (1 + au)^2 / (1 + bv)
$$

(A9)

To calculate the mean and variance of $\hat{N}_{\text{mean}}^*$ from (A9), one must account for any correlation between the numerator and denominator. We have shown that $E$ has an approximate scaled $\chi^2(N_{\text{mean}}^*)$ distribution. In our parameter regime, $N_{\text{mean}}^* \gg 1$, and this distribution becomes approximately normal. Laha and Rohatgi (1979) showed that the standard estimators of the mean and variance of a finite number of independent samples from a normal distribution are uncorrelated. We conclude that $\langle \hat{E} \rangle$ and $\text{var}(E)$, and thus $u$ and $v$, are approximately uncorrelated.

With this result, we calculate the mean and variance of $N_{\text{mean}}^*$, including all terms at least as large as $O(T^{-1})$ (i.e., second order in $a$, $b$). Substituting (A6) and (A7) into (A9), and noting that $\langle u^2 \rangle = \langle v^2 \rangle = 1$ and $\langle u \rangle = \langle v \rangle = \langle uv \rangle = 0$, we obtain (12):

$$
\langle \hat{N}_{\text{mean}}^* \rangle / N_{\text{mean}}^* = \left( (1 + au)^2 / (1 + bv) \right)^{1/2}
$$

$$
= 1 + 2a(u) + a^2(u^2) - b(v)
$$

$$
- 2ab(uv) + b^2(v^2)
$$

$$
= 1 + a^2 + b^2 \approx 1 + \frac{2}{T}.
$$

(A10)

In deriving the final line above, we note that $a^2/b^2 = (2N_{\text{mean}}^*)^{-1} \ll 1$. Similarly,
The matrix $z_i$ in terms of the true normalized principal components to the eigenvalues of matrix. We can apply the same transformation to $l$ eigenvectors, and let $g$ eigenvectors of the true covariance matrix. Let simplify our analysis, we work in the basis of the ei-
and correlation coefficient of these two functions. To two functions are strongly correlated, and this must be small in our parameter regime. Unlike for, these variation of each of these quantities is $O^2$ can be computed. As for, the relative standard de-
matrix), each having a mean and standard deviation that $N$ independent sampling times. From (3) applied to the ei-

$$\delta N_{\text{true}} / N_{\text{true}} = \left( \frac{4}{N_{\text{true}} T} + \frac{2}{T} \right)^{1/2} \approx (2/T)^{1/2}.$$ 

b. Bias and scatter in $\hat{N}_{\text{true}}$

Let $C$ be the true covariance matrix. From (3) applied to the eigenvalues of $C$, we could calculate the true $N^\ast_{12} = (\Sigma, \lambda)^i \Sigma \lambda^j$. Let $C$ be an estimate of $C$ based on $T$ independent sampling times. From (3) applied to the eigenvalues of $C$, we obtain an estimate $\hat{N}^\ast_{12}$ of the true $N^\ast_{12}$. We regard $\hat{N}^\ast_{12}$ as a random variable and estimate its mean and standard deviation. Our estimates assume that the field values observed at each grid point (and hence the true principal components) are normally distributed.

As with $N^\ast_{12}$. $N^\ast_{12}$ is the quotient of two functions (in this case, functions of the eigenvalues of the covariance matrix), each having a mean and standard deviation that can be computed. As for $N^\ast_{12}$, the relative standard deviation of each of these quantities is $O(T^{-1/2})$, which is small in our parameter regime. Unlike for $N^\ast_{12}$, these two functions are strongly correlated, and this must be accounted for in our analysis.

We must first compute the mean, standard deviation, and correlation coefficient of these two functions. To simplify our analysis, we work in the basis of the eigenvectors of the true covariance matrix. Let $E$ be the orthogonal matrix whose columns are these normalized eigenvectors, and let $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, where the $\lambda_i$ are the eigenvalues of $C$. Then

$$\Lambda = E^T C E.$$ 

We can apply the same transformation to $\hat{C}$ to obtain a matrix,

$$\hat{\Lambda} = E^T \hat{C} E.$$ 

The matrix $\hat{\Lambda}$ is no longer exactly diagonal, but it has the same eigenvalues as $\hat{C}$, and its elements can be computed in terms of the true normalized principal components $z_i$: 

$$\hat{\Lambda}_{ij} = \frac{\lambda_i^{1/2} \lambda_j^{1/2}}{T} \sum_{n=1}^T z_i(t_n) z_j(t_n).$$  

(A12)

To compute $\hat{N}^\ast_{12}$, we do not have to find the eigenvalues of $\hat{C}$ or $\hat{L}$. Instead, we can apply (5) to $L$:

$$\hat{N}^\ast_{12} = \frac{(\text{tr} \hat{L})^2}{\text{tr}(L^2)} = \frac{\left( \sum_{i=1}^n \hat{L}_{ii} \right)^2}{\sum_{i,j=1}^n \hat{L}_{ij}}.$$  

(A13)

The numerator $[\text{tr}(\hat{L})]^2 = (\Sigma, \hat{\Lambda})^2$, and the denominator $\text{tr}(L^2) = \Sigma, \hat{\Lambda}^2$.

It is straightforward to show that

$$\hat{\Lambda}_{ij} = \begin{cases} \lambda_i, & j = i \\ 0, & j \neq i \end{cases}$$

and

$$\delta \hat{\Lambda}_{ij} = \lambda_i^{1/2} \lambda_j^{1/2} \begin{cases} (2/T)^{1/2}, & j = i \\ (1/T)^{1/2}, & j \neq i \end{cases}$$

Furthermore, each element of $\hat{L}$ is uncorrelated with all other elements of $\hat{L}$. In the rest of this analysis, we will restrict ourself to the case in which $T$ is large, so each element $\hat{L}_{ii}$, which is a sum of $T$ independent and identically distributed random variables, must have an approximately normal distribution by the central limit theorem. Thus the statistics of $\hat{L}$ can be determined by noting that

$$\hat{L}_{ii} \sim \lambda_i^{1/2} \lambda_i^{1/2} \left[ 1 + (w_i / T)^{1/2} \right] j = i \quad \text{and}$$

$$\hat{L}_{ij} \sim \lambda_i^{1/2} \lambda_j^{1/2} \left[ (w_i / T)^{1/2} \right] j \neq i,$$ 

where the $w_i$ are uncorrelated unit normal random variables and $\sim$ indicates that the two sides have the same PDF.

We are now ready to compute the mean and variance of $\text{tr}(\hat{L})^2$ and $\text{tr}(\hat{L})^2$. To make our results more compact, we define

$$S_m = \sum_{i=1}^N \lambda_i.$$ 

We will also make use of some moments of the unit normal distribution: $\langle w_i \rangle = 1$, $\langle w_i^2 \rangle = 3$, and $\text{var}(w_i^2) = 2$. The statistics of $\text{tr}(\hat{L})$ are as follows:

$$\langle \text{tr}(\hat{L}) \rangle = \sum_{i=1}^N \langle \hat{L}_{ii} \rangle = \sum_{i=1}^N \lambda_i = S_m,$$  

(A14)

$$\text{var}[\text{tr}(\hat{L})] = \sum_{i=1}^N \text{var}[\hat{L}_{ii}] = \sum_{i=1}^N \lambda_i^2 \text{var}[1 + w_i (2/T)^{1/2}]$$

$$= 2 S_m / T,$$

$$\delta[\text{tr}(\hat{L})] = (2 S_m / T)^{1/2}.$$  

(A15)

Note that since each of the $\hat{L}_{ii}$ has a normal distribution, so will $\text{tr}(\hat{L})$. 

The statistics of \( \text{tr}(\hat{L}^2) \) are computed as follows:

\[
\langle \text{tr}(\hat{L}^2) \rangle = \sum_{i=1}^{N} \langle \hat{L}_{ii}^2 \rangle = \sum_{i=1}^{N} \lambda_i^2/T + \sum_{i=1}^{N} \lambda_i^2(1 + w_i/(2/T)^{1/2})^2
\]

\[
= \sum_{i=1}^{N} \lambda_i^2/T + \sum_{i=1}^{N} \lambda_i^2(1 + 2/T)
\]

\[
= S_2 + (S_1^2 + S_2)/T, \quad \text{(A16)}
\]

\[
\text{var}[\text{tr}(\hat{L}^2)] = \sum_{i=1}^{N} \text{var}[\hat{L}_{ii}^2] = \sum_{i=1}^{N} \lambda_i^2 \lambda_i^2 \text{var}(w_i)/T^2 + \sum_{i=1}^{N} \lambda_i^2 \text{var}(1 + w_i/(2/T)^{1/2})^2.
\]

To compute \( \text{var}[1 + w_i/(2/T)^{1/2}] \), we proceed as follows:

\[
\text{var}[1 + w_i/(2/T)^{1/2}] = \langle (1 + w_i/(2/T)^{1/2})^2 \rangle - \langle (1 + w_i/(2/T)^{1/2}) \rangle^2
\]

\[
= 1 + 12(w_i^2)/T + 4(w_i^2)/(2/T)^2 - (1 + 2(w_i^2)/(2/T))^2
\]

\[
= 8/T + 8/T^2.
\]

We deduce that

\[
\text{var}[\text{tr}(\hat{L}^2)] = \sum_{i=1}^{N} \lambda_i^2 \lambda_i^2 /T^2 + \sum_{i=1}^{N} \lambda_i^2 (8/T + 8/T^2)
\]

\[
= 8S_2/T + (S_1^2 + 7S_2)/T^2. \quad \text{(A17)}
\]

\[
\delta[\text{tr}(\hat{L}^2)] = \langle 8S_2/T \rangle^{1/2}, \quad T \gg N_\lambda = S_2^2/S_1. \quad \text{(A18)}
\]

Note that \( N_\lambda \) is analogous to \( N^+_{\eta} \), except with \( \lambda_i^2 \) used in place of \( \lambda_i \). Since this accentuates the relative importance of the highest-variance PCs, \( 1 < N_\lambda < N^+_{\eta} \). Thus it is safe to replace the condition \( T \gg N_\lambda \) by the more conservative condition \( T \gg N^+_{\eta} \), which is the parameter regime assumed throughout this analysis.

The above analysis shows that the dominant contribution to the variability of \( \text{tr}(\hat{L}^2) \) comes from the sum of the diagonal terms \( \hat{L}_{ii}^2 \), and in particular from the sum of the subterms \( w_i(8/T)^{1/2} \) of those diagonal terms. Since \( w_i \) is normally distributed, this sum, and hence \( \text{tr}(\hat{L}^2) \), should also be normally distributed.

Last, the covariance of \( \text{tr}(\hat{L}) \) and \( \text{tr}(\hat{L}^2) \) must be computed. This is greatly simplified by noting that the elements of \( \hat{L} \) are uncorrelated with each other:

\[
\text{cov}[\text{tr}(\hat{L}), \text{tr}(\hat{L}^2)] = \langle [\text{tr}(\hat{L}) - \langle \text{tr}(\hat{L}) \rangle][\text{tr}(\hat{L}^2) - \langle \text{tr}(\hat{L}^2) \rangle] \rangle
\]

\[
= \left( \sum_{i=1}^{N} \lambda_i w_i/(2/T)^{1/2} \right)^2 \text{var}[\hat{L}_{ii}^2]
\]

\[
= \sum_{i=1}^{N} \lambda_i^2 (w_i^2/(2/T)^{1/2})^2 \times [2(w_i/(2/T)^{1/2} + 2(w_i^2 - 1)/T)]
\]

\[
= \frac{4}{T} \sum_{i=1}^{N} \lambda_i^2 w_i^2 = 4S_2/T. \quad \text{(A19)}
\]

The relative standard deviation of \( \text{tr}(\hat{L}) \) is found from (A14) and (A15):

\[
A = \frac{\delta[\text{tr}(\hat{L})]}{\langle \text{tr}(\hat{L}) \rangle} = \frac{(2S_2/T)^{1/2}}{S_1} = \left( \frac{2}{N^+_{\eta}T} \right)^{1/2}. \quad \text{(A20)}
\]

The relative standard deviation of \( \text{tr}(\hat{L}^2) \) is found from (A16) and (A18), neglecting higher-order terms in \( T \):

\[
B = \frac{\delta[\text{tr}(\hat{L}^2)]}{\langle \text{tr}(\hat{L}^2) \rangle} = \frac{(8S_2/T)^{1/2}}{S_2} = \left( \frac{8}{N^+_{\eta}T} \right)^{1/2}. \quad \text{(A21)}
\]

We have argued that \( \text{tr}(\hat{L}) \) and \( \text{tr}(\hat{L}^2) \) are approximately normally distributed. We can normalize them to unit normal random variables \( U \) and \( V \):

\[
\text{tr}(\hat{L}) = (\text{tr}(\hat{L}))(1 + AU), \quad \text{tr}(\hat{L}^2) = (\text{tr}(\hat{L}^2))(1 + BV).
\]

Substituting these expressions into the definition (A13) of \( N^+_{\eta} \),

\[
N^+_{\eta} \approx \left( \text{tr}(L)^2 (1 + AU)^2 \right)^{1/2} \left( \text{tr}(L^2) \right)^{1/2} = \frac{N^+_{\eta}^* (1 + AU)^2}{1 + BV}. \quad \text{(A22)}
\]

Following the analogous derivation of the mean and variance of \( N^+_{\eta} \), we note that \( A \) and \( B \) are both small \([O(T^{-1/2})]\), and we Taylor-expand the second fraction of (A22) up through quadratic terms in \( A \) and \( B \). The new twist is that \( U \) and \( V \) are now strongly correlated. Their correlation coefficient can be computed from (A19), (A15), and (A18):

\[
\langle UV \rangle = \text{cov}[\text{tr}(\hat{L}), \text{tr}(\hat{L}^2)] = \frac{4S_2/T}{(2S_2/T)^{1/2}(8S_2/T)^{1/2}} = \frac{S_1}{(S_2S_4)^{1/2}}. \quad \text{(A23)}
\]

In the limit of either one dominant eigenvalue, or all eigenvalues equal, this correlation \( R \) is unity. For eigenvalues slowly decreasing geometrically like powers of some constant \( \exp(-\mu) \), \( \mu \ll 1 \), \( R \) can be shown to be \((8/9)^{1/2} \approx 0.94\). For most realistic eigenvalue spectra, \( R \) exceeds 0.9, so the numerator and denominator of \( N^+_{\eta} \) are highly correlated.
We now mimic the analysis that led up to Eqs. (A10) and (A11) for the mean and variance of \( \tilde{N}_c \), but retaining terms proportional to \( \langle UV \rangle \). A computation of the mean of \( \tilde{N}_c \) from (A22) leads to (14):

\[
\langle \tilde{N}_c \rangle / \langle N_c \rangle = \frac{1}{1 + \frac{N_c}{T}} \left( \frac{(1 + AU)^2}{1 + BV} \right)
\]

\[
\approx \frac{1}{1 + \frac{N_c}{T}} \left( 1 + 2A(U) + A^2(U^2) - B(V) \right)
\]

\[
- 2AB(UV) + B^2(V^2) \right)
\]

\[
= \frac{1}{1 + \frac{N_c}{T}} \left( 1 + A^2 - 2AB + B^2 \right)
\]

\[
\approx \frac{1}{1 + \frac{N_c}{T}}.
\]

(A24)

In the last step, we have noted that \( A^2, AB, \) and \( B^2 \) are all proportional to \( (N_c T)^{-1} \), which is much smaller than \( N_c / T \) since \( N_c \gg 1 \).

A similar computation of the variance of \( \tilde{N}_c \) shows that

\[
\text{var}[\tilde{N}_c] = \left( \frac{N_c}{1 + \frac{N_c}{T}} \right) \times \left[ \left( \frac{(1 + AU)^2}{1 + BV} \right) - \left( \frac{(1 + AU)^2}{1 + BV} \right) \right]
\]

\[
= \left( \frac{N_c}{1 + \frac{N_c}{T}} \right) \times \left\{ \left( 1 + 4A(U) + 6A^2(U^2) - 2B(V) \right) - \left( 1 + A^2 - 2AB(UV) + B^2(V^2) \right) \right\}
\]

\[
= \left( \frac{N_c}{1 + \frac{N_c}{T}} \right) \times \left\{ (1 + 6A^2 - 8AB + 3B^2) \right\}
\]

\[
- \left( 1 + 2A^2 - 4AB + B^2 \right)
\]

\[
= N_c / (4A^2 - 4AB + B^2). \tag{A25}
\]

(A25)

In the last line, we have neglected the term \( N_c / T \) as being small compared to one, as it is not a leading-order contributor to the variance.

Substituting for \( A, B, \) and \( R \) in (A25) using (A20), (A21), and (A23), and converting from variance to standard deviation, we obtain (15):

\[
\delta \tilde{N}_c / \langle N_c \rangle = cT^{-1/2},
\]

where

\[
c^2 = 8 \left( \frac{1}{N_c^2} + \frac{1}{N_s} - 2R \left( \frac{1}{N_s} \right) \right). \tag{A26}
\]

(A26)

Because \( R \) is close to 1, and \( N_s \) can be nearly as large as \( N_c^2 \), the last term in \( c^2 \) partially cancels the other two, and the relative standard deviation of \( \tilde{N}_c \) is quite small.

In fact, if all eigenvalues are equal, or if one mode explains all the variance, \( c \) is exactly zero, the scatter to \( O(T^{-1/2}) \) is negligible, and the relative standard deviation must be of higher order in \( T \).

REFERENCES


