**Nonlinear Processes
in Geophysics**

Statistical inference from atmospheric time series: detecting trends and coherent structures

A. Gluhovsky

Department of Earth and Atmospheric Sciences and Department of Statistics, Purdue University,
West Lafayette, IN 47907, USA

Abstract. Standard statistical methods involve strong assumptions that are rarely met in real data, whereas resampling methods permit obtaining valid inference without making questionable assumptions about the data generating mechanism. Among these methods, subsampling works under the weakest assumptions, which makes it particularly applicable for atmospheric and climate data analyses. In the paper, two problems are addressed using subsampling: (1) the construction of simultaneous confidence bands for the unknown trend in a time series that can be modeled as a sum of two components: deterministic (trend) and stochastic (stationary process, not necessarily an i.i.d. noise or a linear process), and (2) the construction of confidence intervals for the skewness of a nonlinear time series. Non-zero skewness is attributed to the occurrence of coherent structures in turbulent flows, whereas commonly employed linear time series models imply zero skewness.

1 Introduction

With the availability of new sources of data, time series analysis has been playing an ever-increasing role in atmospheric and climate studies. The problem is that conventional statistical methods are "based on certain probabilistic assumptions about the nature of the physical process that generates the time series of interest. Such mathematical assumptions are rarely, if ever, met in practice" (Ghil et al., 2002). One common assumption is that observations are normally distributed. Yet in reality, distributions are often not normal, such as those for the velocity field in a turbulent flow (Lesieur, 2008), the precipitation amount, or the economic damage from extreme weather events (Katz, 2002; Katz et al., 2002), and new advances in statistics have made it clear

that even slight departures from normality can be a source of concern (e.g., Wilcox, 2003). Another questionable assumption is that of a linear model for the observed time series, whereas the real data generating mechanism (DGM) is inherently nonlinear, so that estimation commonly based on a fitted linear models may be misleading (e.g., Gluhovsky, 2008).

Meanwhile, recent progress in computer-intensive (aka bootstrap or resampling) methods makes it possible to avoid reliance on questionable assumptions in time series analysis. One such method, the subsampling (Politis et al., 1999), is particularly suitable for atmospheric and climate time series. In this paper, subsampling techniques are suggested to address two fundamental problems in atmospheric and climate dynamics: trends and coherent structures (CSs) that are described in Sect. 2. As Phillips (2005) noted, "no one understands trends, but everyone sees them in the data", and in spite of observational successes, the problem of describing turbulent flows with CSs remains a formidable theoretical challenge (Tabeling, 2002). In Sect. 3, subsampling is briefly introduced and, as an example, the construction of subsampling confidence intervals for the skewness of an observed time series is provided (positive skewness indicates the presence of CSs). Then the subsampling technique developed for confidence intervals is extended to that for the construction of simultaneous confidence bands for unknown trends.

2 Motivating problems

2.1 Trends in climate variables

As with other important concepts, such as turbulence or coherent structures, there is no commonly accepted definition of a trend, and though even today the trend remains relatively little understood (e.g., White and Granger, 2011), it is usually taken as a smooth function representing systematic developments.

One difficulty in trend analysis is that what one perceives as a deterministic trend may well be produced by a purely stochastic mechanism (as with random walks or long-memory processes); such an artifact is sometimes called a stochastic trend. In practice, when a relatively short portion of the series is only available, the two possibilities (deterministic or stochastic trend) are often indistinguishable statistically (e.g., Fatichi et al., 2009), though sometimes conclusive statistical evidence can be obtained. For the temperature data, for instance, Beran and Feng (2002) found statistical evidence for a deterministic trend by fitting their SEMIFAR model, Rybski and Bunde (2009) detected linear trends using detrended fluctuation analysis. And, to add more confusion, non-stationarity in the mean may also cause the long-memory effect (Bhattacharya et al., 1983).

Therefore, for a number of reasons (see also Ashley and Patterson, 2010; Mudelsee, 2010), it is often practical to follow the classical approach to model the time series as a sum,

$$Z_t = m_t + \epsilon_t, \qquad (1)$$

of an unknown function, trend m_t, and a stationary zero-mean process ϵ_t. Then the trend may be estimated from the data by nonparametric techniques such as local polynomial regression with global (Cleveland and Devlin, 2009; Efromovich, 1999) or local (Ruppert, 1997; Gluhovsky and Agee, 2007) bandwidth selection or by using wavelets (Craigmile et al., 2004). For example, the red curve in Fig. 1 shows a trend computed via local polynomial regression in the global temperature series that plays a prominent role in climate change studies (the gray curve, from Cowpertwait and Metcalfe, 2009). The series of 1800 data points gives global monthly anomalies (°C) from January 1856 to December 2005, relative to the average of the period 1961–1990.

Another difficulty in assessing trends, as well as time series parameters, is that estimates without any measure of their validity are not very useful. In nonlinear regression, when ϵ_t in Eq. (1) are independent identically distributed random variables, this is provided by the *simultaneous confidence bands* (SCBs, e.g., Eubank and Speckman, 1993) that quantify the associated uncertainty (similar to *confidence intervals* (CIs) in classical statistics), so that the two functions, l_t and u_t, in Eq.(2)

$$P\left(l_t < m_t < u_t\right) \approx 0.90, \qquad (2)$$

trap the unknown trend with a given confidence, say, 90 %.

With dependence present (i.e., when ϵ_t is a more general stationary process, which is typically the case in geosciences), constructing SCBs becomes a considerably more difficult problem, which has been addressed in various ways (Bühlmann, 1998; Wu and Zhao, 2007; Mudelsee, 2010). When handled via subsampling, however, the problem becomes similar (as will be seen in Sect. 3.2) to a more familiar one of obtaining subsampling CIs for parameters of a time series (described in Sect. 3.1). The red and the blue dashed
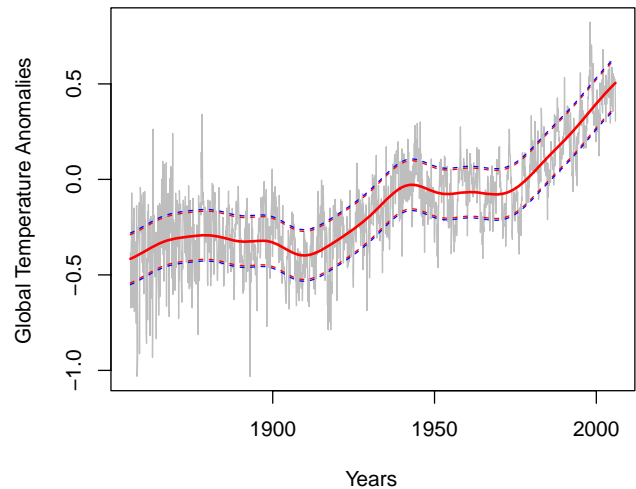

Fig. 1. Global monthly temperature anomalies (°C), relative to 1961–1990 (gray curve) with a trend estimate (thick red curve), a 90 % SCB for the trend (dashed red curves) and calibrated 90 % SCB for the trend (dashed blue curves).

curves in Fig. 1 indicate the two versions of 90 % subsampling SCBs for the trend in the temperature series obtained in Sect. 3.2.

2.2 Coherent structures in turbulent flows

The study of organized structures in turbulent flows began with the well-known laboratory experiments by Brown and Roshko (1974) who provided the first documented visual evidence that the mixing layer can be dominated by what are now called *coherent structures* (CSs). The definition of CSs continues to be somewhat evasive, but they are commonly viewed as organized long-lived motions that spontaneously arise, trap much energy, and cover the whole spectrum of fluid motions (down to the Kolmogorov scale). Perhaps the most relevant atmospheric examples can be found in organized structures that occur on a variety of spatial scales within convective boundary layers that form during cold air outbreaks over warmer water (Agee, 1987).

Meanwhile, turbulent flows with CSs are characterized by non-Gaussian statistics (e.g., Wyngaard and Weil, 1991; Maurizi, 2006) reflected in the values of skewness and kurtosis that are different from those for a normal distribution (0 and 3, respectively). Moreover, in numerical simulations (e.g., Farge et al., 2003; Ruppert-Felsot et al., 2005), turbulent flows were separated into coherent and incoherent components using wavelet transforms. The coherent part, represented by only a small fraction of the wavelet coefficients, retained the total flow dynamics and statistical properties, notably non-Gaussian skewness and kurtosis, while the incoherent part was characterized by Gaussian statistics. Thus, there is considerable interest in accurate estimation
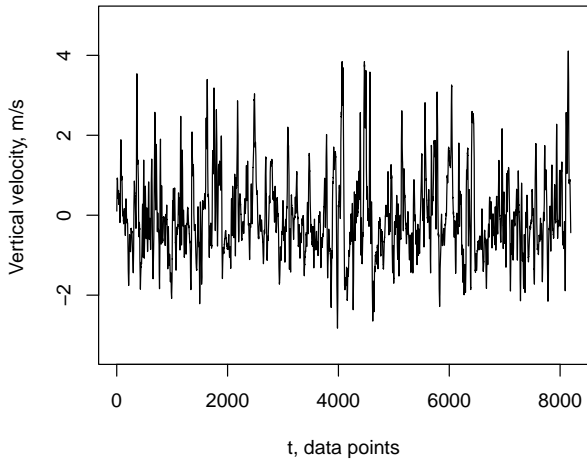
Fig. 2. Record of 20-Hz aircraft vertical velocity measurements over Lake Michigan.

of skewness and kurtosis from time series records, however, common practices may produce misleading results.

As a typical example, consider a time series of the vertical velocity of wind in a convective boundary layer during an outbreak of a polar air mass over the Great Lakes region. The record in Fig. 2 (that has passed a test for stationarity Gluhovsky and Agee, 1994) consists of 8192 data points over about 29 km across Lake Michigan, 50 m above the lake. The sample mean, variance, skewness, and kurtosis of the vertical velocity computed from the record are -0.04, 1.06, 0.83, and 4.10, respectively. Although large skewness and kurtosis may result from nonlinearities in the underlying data generating mechanism (DGM), sample characteristics like these (routinely obtained in field programs as well as in laboratory experiments and computer simulations) are just point estimates (our "best guesses") of the true values of the parameters. Therefore, to learn how far one can trust these numbers, CIs are employed.

A 90 % CI is the range of numbers that traps the unknown parameter with probability 0.90 called the *coverage probability*. Also referred to as a *nominal* or *target* coverage probability (e.g., Davison and Hinkley, 1997), it is attained only if the assumptions underlying the method for the CI construction are met. Since in geosciences this is rarely the case, the *actual* coverage probability may differ from the target level (sometimes considerably). For example, when the DGM (or a model time series) is linear, CIs for the mean or the variance of the time series may be found analytically, but the common practice of computing CIs from fitted linear models may result in erroneous CIs when the real DGM is, in fact, nonlinear (Gluhovsky and Agee, 2007; Gluhovsky, 2008). Meanwile, CIs for the skewness cannot be based on linear models, which imply zero skewness. This is where subsampling becomes instrumental since subsampling CIs do not rely on questionable assumptions.

3 Subsampling confidence sets

3.1 Subsampling confidence intervals for parameters of time series

Although CIs for parameters of a *model* time series are of no particular interest, they can be obtained via Monte Carlo simulations with the model (which leads to the idea of subsampling). To construct, say, a 90 % CI for a (known) time series parameter θ, one could generate lots of the time series realizations and compute from them estimates of, for example, $q_{0.05}$ and $q_{0.95}$, the 0.05 and 0.95 quantiles of the distribution of the *root*, $\hat{\theta} - \theta$. Then an *equal-tailed* (Politis, 1998) 90 % CI for θ is

$$(\hat{\theta} - q_{0.95}, \hat{\theta} - q_{0.05}), \tag{3}$$

where $\hat{\theta}$ is a sample estimate of θ.

Or, one may compute $Q_{0.90}$, a 0.90 quantile of the distribution of $|\theta - \hat{\theta}|$, i.e., $\mathrm{Prob}(|\theta - \hat{\theta}| < Q_{0.90}) = 0.90$, then a *symmetric* (Hall, 1988) 90 % CI for θ is given by

$$(\hat{\theta} - Q_{0.90}, \hat{\theta} + Q_{0.90}). \tag{4}$$

In real life situations, where the DGM (the model) is unknown and typically only one record is available, subsampling comes to the rescue by replacing independent computer generated realizations from the known model by *subsamples*, or blocks of consecutive observations from the single record that retain the dependence structure of the time series (Politis et al., 1999). Underscored below are the first, intermediate, and the last block, all of the same length b (the *block size*) in a record of a time series, Y_t, containing n observations and, therefore, $n - b + 1$ blocks:

$$\underbrace{Y_1, ..., Y_b}_{b}, \; ..., \; \underbrace{Y_i, ..., Y_{i+b-1}}_{b}, \; ..., \; \underbrace{Y_{n-b+1}, ..., Y_n}_{b}. \tag{5}$$

Subsampling does not require that any model, linear or nonlinear, be fitted to the data, and it works in complex dependent data situations under the weakest assumptions among other computer-intensive techniques. Still, models are useful since it is with models that one may assess the *actual* coverage of CIs via Monte Carlo (MC) simulations: with known model, one may generate numerous records, compute from each one the subsampling CI, and estimate its coverage probability by counting the fraction of times the known parameter value, θ, was within the CI.

With this in mind, consider the model (Lenschow et al., 1994),

$$Y_t = X_t + a(X_t^2 - 1), \tag{6}$$

where X_t is a first order autoregressive process (AR(1)),

$$X_t = \phi X_{t-1} + \epsilon_t, \tag{7}$$

$0 < \phi < 1$ and a are constants, and ϵ_t is a white noise process (a sequence of uncorrelated random variables with mean 0 and variance σ_ϵ^2) with $\sigma_\epsilon^2 = 1 - \phi^2$ so that $\sigma_X^2 = 1$.
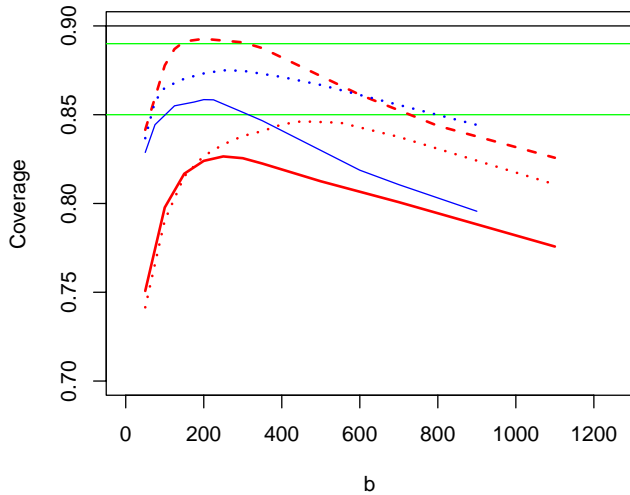
Fig. 3. Actual coverage probabilities of 90 % subsampling CIs for the skewness (red) and variance (blue) of nonlinear time series (6) at $a = 0.145$: $n = 2048$ (solid), $n = 4096$ (dotted), calibrated CIs at $n = 2048$ (dashed). Horizontal green lines denote 0.85 and 0.89 levels.

AR(1) with a Gaussian white noise is widely employed in studies of climate as a default model for correlated time series (e.g., von Storch and Zwiers, 1999; Percival et al., 2004). When the white noise in model (7) is not Gaussian, the model may exhibit nonlinear behavior and is referred to as an *implicit* nonlinear model (Fan and Yao, 2003), as opposed to *explicit* nonlinear model (6), where AR(1) is altered with a nonlinear component.

At $a = 0.145$, the mean, variance, skewness, and kurtosis of Y_t (in model 6), are respectively,

$$M = 0, \quad V = 1 + 2a^2 \approx 1.04, \quad S = (6a + 8a^3)/V^{1.5}) \approx 0.84,$$
$$K = (3 + 60a^2 + 60a^4)/V^2 \approx 3.95,$$

i.e., they are close to the corresponding sample characteristics $(-.04, 1.06, 0.83, 4.10)$ of the vertical velocity time series in Fig. 2. Thus, model (6) might provide a better description for that series than linear models, which inherently have zero skewness. In simulations below, $\phi = 0.67$ and the the records contain 2048 data points, which permits to imitate the dependence structure of the vertical velocity time series as characterized by autocorrelation functions.

Figure 3 presents the actual coverage probabilities of subsampling CIs for the variance (blue curves) and the skewness (red curves) computed via MC simulations with model (6) at $a = 0.145$. Subsampling CIs were found following Eq. (4) with $\hat{\theta}$ now being the sample skewness (or variance) from the whole record, and quantile $Q_{0.90}$ being estimated from all values of $|\hat{\theta} - \hat{\theta}_i|$, where $\hat{\theta}_i$ was the sample skewness (or variance) computed from the i-th subsample (5).

The curve for the variance in Fig. 3 (solid blue) corresponding to $n = 2048$ has the maximum of about 0.87 at

$b = 200$ (exceeding by far actual coverages of CIs based on linear models e.g., Gluhovsky, 2008), while the curve for the skewness (solid red), where conventional CIs are unavailable, provides the maximum of about 0.83 at $b = 250$. Estimating the skewness does require longer records than those for the variance (e.g., Gluhovsky and Agee, 1994; Lenschow et al., 1994).

A simple way to improve the coverage is to increase (when feasible) the record length. The dotted curves in Fig. 3 show the increased coverage probabilities for the variance and skewness (blue and red, respectively) due to records of 4096 observations. Otherwise, the so-called *calibration* can be used (Politis et al., 1999): nominal 0.90 CIs in the subsampling procedure are replaced with those of higher confidence level, which may be determined via MC simulations with an approximating model (such as model (6) at $a = 0.145$ for the time series in Fig. 2). For example, the dashed red curve in Fig. 3 was obtained for the skewness in case $n = 2048$ by replacing 0.90 subsampling CIs (which resulted in the solid red curve) with 0.96 subsampling CIs. The dashed curve demonstrates that coverage probabilities close to the target can be achieved for a range of block sizes (the curve is above 0.89 level here at $b \in (125, 350)$). A similar curve for the variance (not shown) that also approaches the target (0.90 at $b = 150$) was obtained by replacing 0.90 subsampling CIs (which resulted in the solid blue curve) with 0.94 subsampling CIs. This also brings the problem of the block size choice in subsampling addressed in Sect. 3.3.

It was found that a 90 % subsampling CI (Eq. 4) for the skewness of the observed time series in Fig. 2 is $(0.63, 1.02)$, whereas a calibrated one with 0.90 coverage (i.e., a 96 % subsampling CI (Eq. 4), according to the approximating model) is $(0.41, 1.24)$. Although the calibrated one is larger, both serve the purpose of confirming that the skewness of the vertical velocity time series is positive, thus indicating the presence of CSs in the flow.

3.2 Subsampling confidence bands for trends

To construct subsampling SCBs for the unknown trend function in Eq. (1), the subsampling procedure of Sect. 3.1 had to be modified to include sample estimates of the trend in place of those for the skewness. Therefore, functions l_t and u_t in Eq. (2) in this case were computed as

$$l_t = \hat{m}_t - Q_{0.90}, \quad u_t = \hat{m}_t + Q_{0.90}, \tag{8}$$

where $\hat{m}_t$ was the sample estimate of the trend computed (via local linear regression) from the whole record, and $Q_{0.90}$ was the 0.90 quantile of the distribution of the maximum of the absolute value of the trend in residuals $Z_t - \hat{m}_t$, estimated via subsampling. That is, the statistic of interest here, $|Z_t - \hat{m}_t|$, was evaluated over subsamples of the record, instead of independent realizations of the time series in case of MC simulations (cf., Eq. 4).
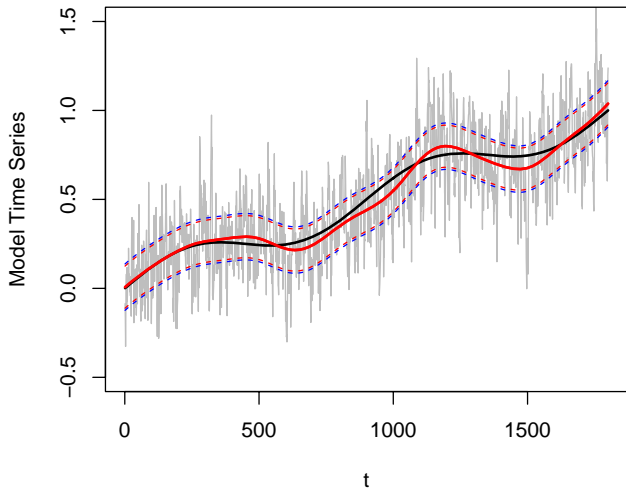
Fig. 4. A realization of time series (9) at $a = 0$ (gray curve), trend function m_t (thick black curve), a trend estimate computed from the realization (thick red curve), a 90 % SCB for the trend (dashed red curves) and calibrated 90 % SCB for the trend (dashed blue curves).

Fig. 5. Actual coverage probabilities of 90 % subsampling SCBs for the trend in model (9) at $a = 0$ (solid red curve), $a = 0.2$ (blue curve), and for a calibrated SCB at $a = 0$ (dotted red curve). Horizontal green lines denote 0.85 and 0.89 levels.

To get an idea of the method's performance, the coverage of the thus obtained SCBs was assessed, as before, by MC simulations with a model time series (representing that in Fig. 1),

$$Z_t = m_t + Y_t. \tag{9}$$

In model (9), the trend function, $m_t = 0.1\sin(4\pi t/n) + t/n$ (the black curve in Fig. 4), is added to the stochastic process (6) considered in Sect. 3.1, which plays the role of ϵ_t in Eq. (1). The gray curve in Fig. 4 shows the total signal Z_t, an example of estimated trend is presented by the red curve, and the dashed red curves show a 90 % subsampling SCB.

For simulations with model (9), records of 1800 observations were used, as in the temperature series in Fig. 1, whose correlations in residuals are roughly described by the autocorrelation function of time series (7) with $\phi = 0.67$. Figure 5 demonstrates the actual coverage of 90 % subsampling SCBs for the trend in series (9) found, again, via MC simulations. The solid red curve shows the actual coverage (at various block sizes) when the noise is linear ($a = 0$ in Eq. (6), the blue curve corresponds to the nonlinear noise with $a = 0.2$. It can be seen that nonlinearity practically does not affect the coverage, but both SCBs somewhat undercover, reaching the maximum actual coverage of slightly over 0.84 at $b = 600$. Hence a calibration might be in order, and the dotted red line in Fig. 5 corresponds to the calibrated subsampling SCB at $a = 0$ (nominal 94 % SCB that provides the actual coverage of 0.90). It turns out, however, that the nominal and calibrated SCBs here (shown by the dashed red and blue curves, respectively, in Fig. 4) are very close, so that the calibration in this case is probably unnecessary.

In practice, one may easily compare subsampling SCBs (or CIs) with calibrated ones at various confidence levels and
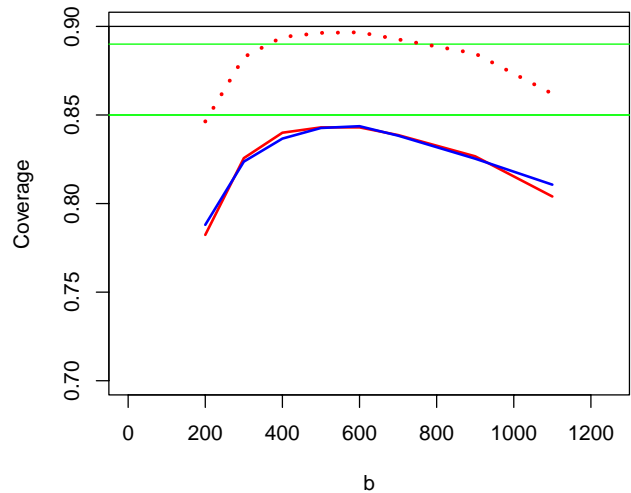
(before looking for and running an approximating model) decide, depending on the purpose of the study, whether calibration is justified. For example, subsampling SCB for the temperature time series in Fig. 1 does not seem to need a calibration (see the last paragraph of the Sect. 3.3).

3.3 Choice of the block size

As seen in Figs. 3 and 5, coverage probabilities of subsampling CIs and SCBs depend considerably on the block size b, and the numbers characterizing the maximum actual coverage in previous sections were found using the optimal blocks ($b = 200$ for the variance and $b = 250$ for the skewness in Sect. 3.1, or $b = 600$ for the trend in Sect. 3.2) obtained via MC simulations with models. With only one record available in practice, however, the choice of the block size becomes the most difficult problem in subsampling (shared by all blocking methods). The asymptotic result (Politis et al., 1999),

$$b \to \infty \quad \text{and} \quad b/n \to 0 \quad \text{as} \quad n \to \infty, \tag{10}$$

that the block size needs to tend to infinity with the sample size but slower, does not help to choose the block size for relatively short atmospheric and climatic records.

To handle this problem, one more resampling technique has been developed (Gluhovsky et al., 2005) for computing the optimal block size from one record in case of subsampling CIs (Eq. 3). Recall (see Sect. 3.1) that to assess the actual coverage of subsampling CIs via MC simulations, one generates numerous records, computes from each one the subsampling CI, then estimates its coverage probability by counting the fraction of times the known parameter was within the CI, and chooses optimal b as that providing the

best coverage. Now, with only one record available, independent realizations generated from a model are replaced by *pseudo realizations* obtained from the single record via the following procedure (a version of the circular bootstrap Politis and Romano, 1992). The record of n data points is "wrapped" around a circle, then $p < n$ points on the circle are chosen at random (following a uniform distribution on the circle) as starting points for p consecutive segments of a pseudo realization. Thus the length of each segment is n/p (n should be a multiple of p) and the pseudo realization has length n. The procedure is repeated to generate numerous pseudo realizations (from the same record) that substitute independent realizations generated from a model time series. Then the "coverage" is determined as before, but the *correct* coverage proves very different. This is because the maxima of the curves analogous to any of those in Figs. 3 and 5 (but based on pseudo realizations) vary wildly, depending on the initial record used to generate the pseudo realizations. It turns out, however, that such curves essentially retain the shape of the corresponding ones obtained via MC simulations, thus indicating a suitable block size to be used in subsampling.

The procedure was successfully used to construct subsampling CIs for the mean and variance (Gluhovsky et al., 2005), skewness (Gluhovsky, 2008), and kurtosis (Gluhovsky and Agee, 2009). Employed in all these studies were CIs (Eq. 3), where the coverage probability could be estimated as the fraction of times the sample estimate (used in place of the parameter known in MC simulations) was within the CI, and the block size choice was determined based on coverage.

In CI (Eq. 4), however, the sample estimate of the parameter is *always* within the CI. In this work, therefore, it was found that dependence of $Q_{0.90}$ on block size b could be used in place of that for the coverage in block size selection. In Fig. 6, for example, each blue dashed curve is based on pseudo realizations obtained from two different records ($n = 1800$, $p = 30$) generated from model (9) (one record for each curve). By comparing the dashed curves with those in Fig. 5, one can see that the they indicate a range of block sizes ($b \in (500, 800)$) appropriate for subsampling. Return now to a real life example: the global temperature series in Sect. 3.1. The solid red curve in Fig. 6 was computed from the time series in Fig. 1 in the same way as the dashed ones from model (9). Then $b = 600$ was chosen as a suitable block sizes and subsampling with $b = 600$ was carried out, resulting in the SCB shown by the red dashed curves in Fig. 1. Calibration was then applied (based on 94 % SCB, according to approximating model (9) resulting in the SCB shown by the blue dashed curves. Similar to Fig. 4, the difference between the original and calibrated SCBs in this case is negligible.
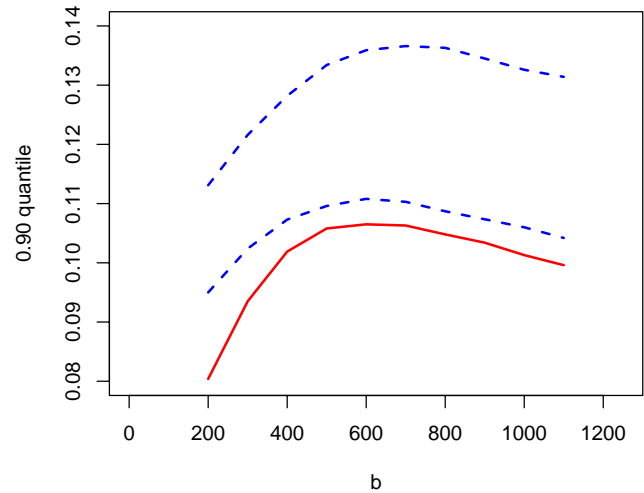

Fig. 6. Quantiles $Q_{0.90}$ in Eq. (8) estimated for different b using pseudo realizations generated from one record: each blue dashed curve is based on pseudo realizations generated from a record of model time series (9), while the solid red curve is based on pseudo realizations generated from the record in Fig. 1.

4 Conclusions

The purpose of this simulation study was to demonstrate that subsampling techniques may be developed to obtain valid statistical inference in a variety of problems, where traditional time series analyses are hindered due to nonlinear data generating mechanisms and limited records. Trends in climate variables and coherent structures (CSs) in turbulent flows are two important problems of this kind considered in the paper. Subsampling simultaneous confidence band (SCB) for the trend in the time series in Fig. 1 confirms the possibility of an increasing temperature trend, and subsampling confidence interval (CI) for the skewness of time series in Fig. 2 confirms that the vertical velocity skewness is positive, thus suggesting the presence of CSs in the flow - inference unattainable by conventional statistical methods. Other topics of intense development, where applying subsampling should be advantageous, are extreme events and long-range dependence (e.g., Nordman and Lahiri, 2005; Rust, 2009; Jach et al., 2011), i.e., problems arising in analysis of heavy-tailed and/or long-memory time series, where common CIs based on asymptotic maximum likelihood fail to capture the real variability (Kallache et al., 2005).

The new tool for the analysis of nonlinear time series presented here is the construction of subsampling SCBs for trends. Our previous work on subsampling (including a resampling technique for selecting the optimal block size for subsampling CIs – the most difficult practical problem in subsampling shared by all blocking methods) focused on *equal-tailed* CIs for parameters of nonlinear time series, the skewness and kurtosis in particular, as they are closely

related to CSs. Because subsampling SCBs for trends are developed in this paper as extensions of *symmetric* CIs (Eq. 4), modifications in subsampling procedures proved necessary, especially in the block size selection. The new implementation of the subsampling procedure that permits a straightforward extension to trends has also provided an opportunity to update the previous analyses of the vertical velocity time series with that based on symmetric subsampling CIs and a longer record (of 8192 data points vs the record of 4096 used previously), and to make this paper self-contained. Also, although almost all published work on bootstrap CIs has focused on equal-tailed intervals, symmetric CIs are often shorter and have better coverage accuracy (Hall, 1988). Thus, it seems useful to have subsampling techniques developed for the construction of both.

Time series models with statistical properties similar to those of observed time series were extensively used in the paper to evaluate (via Monte Carlo simulations) the performance of subsampling techniques. Such approximating models are also required to determine the level of calibration. However, calibration may not be necessary, which can be determined (*before* looking for and running an approximating model) by comparing subsampling SCBs or CIs with the calibrated ones at different confidence levels. Calibration will only be justified if the difference between the original and calibrated confidence sets will be considered significant for the problem at hand. For example, as noted in the end of Sect. 3.3, subsampling SCB for the temperature time series in Fig. 1 does not need calibration.

Under very weak assumptions, the subsampling methodology provides large-sample CIs of asymptotically correct coverage (Politis et al., 1999), but developing practical subsampling procedures for records of limited length involves the block size selection and, when calibration is required, the search for an approximating model (such as model (6) for the time series in Fig. 2). Like the former, the latter problem may also be difficult since linear models are inappropriate, whereas the multitude of nonlinear models is overwhelming. As way to handle the problem, one might consider physically sound low-order models that possess fundamental properties of fluid dynamical equations (in the spirit of Lorenz, 1963, 1982 and Obukhov, 1973, see also Gluhovsky, 2006), thereby infusing more physics into time series analysis. From the perspective of complexity theory, the development of appropriate statistical methods for atmospheric and climate data analysis should be based on time series spawned by the underlying dynamics rather than on traditional time series models (cf., Nicolis and Nicolis, 2007). Among other problems that may benefit from the physical insight are those arising in estimating the trend function, such as the choice of the bandwidth in local polynomial regression, which currently is based entirely on statistical considerations (e.g., Cleveland and Devlin, 2009; Efromovich, 1999; Gluhovsky and Gluhovsky, 2007). The problem is very relevant but beyond the scope of this paper devoted to finding the

ways to decide how much importance is reasonable to confer on estimated trends.

Finally, the evolution and subsequent growth of CSs may represent an underlying physical mechanism that can lead to extreme events. Due to the incidence of CSs in turbulent flows (indicated by non-Gaussian velocity skewness and kurtosis), the tails of probability density functions become heavy, thus increasing the probability of extremes (McWilliams, 2007), and long-term trends observed in meteorological variables may alter conditions for the formation of CSs, thus affecting the intensity and frequency of extreme events.

Edited by: R. Donner
Reviewed by: three anonymous referees

References

Agee, E. M.: Mesoscale cellular convection over the oceans, Dyn. Atmos. Oceans, 10, 317–341, 1987.

Ashley, R. A. and Patterson, D. M.: Apparent long memory in time series as an artifact of a time-varying mean: considering alternatives to the fractionally integrated model, Macroecon. Dyn., 14, 59-87, 2010.

Beran, J. and Feng, Y.: SEMIFAR models - a semiparametric approach to modelling trends, long-range dependence and nonstationarity, Comput. Stat. Data An., 40, 393–419, 2002.

Bhattacharya, R. N., Gupta, V. K., and Waymire, E.: The Hurst effect under trends, J. Appl. Prob. 20, 649-662, 1983.

Brown, G. L. and Roshko, A.: On density effects and large structure in turbulent mixing layers, J. Fluid Mech., 64, 775–816, 1974.

Bühlmann, P.: Sieve bootstrap for smoothing in nonstationary time series, Ann. Statist., 26, 48-83, 1998.

Cleveland, W. S. and Devlin, S. J.: Locally-Weighted Regression: an Approach to Regression Analysis by Local Fitting, J. Amer. Stat. Assoc., 83, 597–610, 1988.

Cowpertwait, P. W. S. and Metcalfe, A. V.: Introductory Time Series with R, Springer, 2009.

Craigmile, P. F., Guttorp, P., and Percival, D. B.: Trend assessment in a long memory dependence model using the discrete wavelet transform, Environmetrics, 15, 35-313, 2004.

Davison, A. and Hinkley, D.: Bootstrap Methods and Their Application, Cambridge University Press, 1997.

Efromovich, S.: Nonparametric Curve Estimation, Springer, 1999.

Eubank, R. L. and Speckman, P. L.: Confidence bands in nonparametric regression, J. Amer. Stat. Assoc., 88, 1287–1301, 1993.

Fan, J. and Yao, Q.: Nonlinear Time Series, Springer, 2003.

Farge, M., Schneider, K., Pellegrino, G., Wray, A. A., and Rogallo, R.: Coherent vortex extraction in three-dimensional homogeneous turbulence: Comparison between cvs-wavelet and pod-fourier decompositions, Phys. Fluids, 15, 2886-2896, 2003.

Fatichi, S., Barbosa, S. M., Caporali, M., and Silva, M. E.: Deterministic versus stochastic trends: Detection and challenges, J. Geophys. Res., 114, D18121, doi:10.1029/2009JD011960, 2009.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K, Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A, Tian, Y., Varadi, F., and Yiou, P.: Advanced spectral methods for climate time series, Rev. Geophys., 40(1), 1–41, 2002.

Gluhovsky, A.: Energy-conserving and Hamiltonian low-order models in geophysical fluid dynamics, Nonlin. Processes Geophys., 13, 125–133, doi:10.5194/npg-13-125-2006, 2006.

Gluhovsky, A.: Subsampling methodology for the analysis of nonlinear atmospheric time series, in: Nonlinear Time Series Analysis in the Geosciences, edited by: Donner, R. V. and Barbosa, S. M., Springer, 3–16, 2008.

Gluhovsky, A. and Agee, E.: A definitive approach to turbulence statistical studies in planetary boundary layers, J. Atmos. Sci., 51, 1682–1690, 1994.

Gluhovsky, A. and Agee, E. M.: On the analysis of atmospheric and climatic time series, J. Appl. Meteorol. Climatol., 46, 1125–1129, 2007.

Gluhovsky, A. and Agee, E. M.: Estimating higher-order moments from time series observed in convective boundary layers, J. Appl. Meteorol. Climatol., 48, 1948–1954, 2009.

Gluhovsky, I. and Gluhovsky, A.: Smooth location dependent bandwidth selection for local polynomial regression, J. Amer. Stat. Assoc., 102, 718–725, 2007.

Gluhovsky, A., Zihlbauer, M., and Politis, D. N.: Subsampling confidence intervals for parameters of atmospheric time series: block size choice and calibration, J. Stat. Comput. Simul., 75, 381–389, 2005.

Hall, P.: On symmetric bootstrap confidence intervals, J. Royal Stat. Soc. B, 50, 35–45, 1988.

Jach, A., McElroy, T., and Politis, D. N.: Subsampling inference for the mean of heavy-tailed long-memory time series, J. Time Ser. Anal., doi:10.1111/j.1467-9892.2011.00742.x, in press, 2011.

Kallache, M., Rust, H. W., and Kropp, J.: Trend assessment: applications for hydrology and climate research, Nonlin. Processes Geophys., 12, 201–210, doi:10.5194/npg-12-201-2005, 2005.

Katz, R. W.: Techniques for estimating uncertainty in climate change scenarios and impact studies, Climate Res., 20, 167-185, 2002.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Res., 25, 1287-1304, 2002.

Lenschow, D., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, J. Atmos. Oceanic Technol., 11, 661-673, 1994.

Lesieur, M.: Turbulence in Fluids, Springer, 2008.

Lorenz, E. N.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130-141, 1963.

Lorenz, E. N.: Low-order models of atmospheric circulations, J. Meteorol. Soc. Jpn., 60, 255-267, 1982.

Mudelsee, M.: Climate Time Series Analysis, Springer, 2010.

Maurizi, A.: On the dependence of third- and fourth-order moments on stability in the turbulent boundary layer, Nonlin. Processes Geophys., 13, 119–123, doi:10.5194/npg-13-119-2006, 2006.

McWilliams, J. C.: Extreme events in geophysical turbulence and waves: What populates the tails of distribution functions?, in: Extreme Events, University of Hawaii at Manoa, 23–26, 2007.

Nicolis, G. and Nicolis, C.: Foundations of Complex Systems, World Scientific, 2007.

Nordman, D. J. and Lahiri, S. N.: Validity of the sampling window method for long-range dependent linear processes, Economet. Theor., 21, 1087–1111, 2005.

Obukhov, A. M.: On the problem of nonlinear interactions in fluid dynamics, Gerlands Beitr. Geophys., 82, 282-290, 1973.

Percival, D., Overland, J., and Mofjeld, H.: Modeling North Pacific climate time series, in: Time Series Analysis and Applications to Geophysical Systems, edited by: Brillinger, D., Robinson, E., and Schoenberg, F., Springer, 151-167, 2004.

Phillips, P. C. B.: Challenges of Trending Time Series Econometrics, Math. Comput. Simulat., 68, 401–416, 2005.

Politis, D. N.: Computer-intensive methods in statistical analysis, IEEE Signal Proc. Magazine, 15, 39–55, 1998.

Politis, D. N. and Romano, J. P.: A circular block-resampling procedure for stationary data, in: Exploring the Limits of Bootstrap, edited by: LePage, R. and Billard, L., John Wiley and Sons, 263-270, 1992.

Politis, D. N., Romano, J. P., and Wolf, M.: Subsampling, Springer, 1999.

Ruppert, D.: Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation, J. Amer. Stat. Assoc., 92, 1049–1062, 1997.

Ruppert-Felsot, J., Praud, O., Sharon, E., and Swinney, H. L.: Extraction of coherent structures in a rotating turbulent flow experiment, Phys. Rev. E, 72, 016311, doi:10.1103/PhysRevE.72.016311, 2005.

Rust, H. W.: The effect of long-range dependence on modelling extremes with the generalised extreme value distribution, Eur. Phys. J. Special Topics, 174, 91–97, 2009.

Rybski, D. and Bunde, A.: On the detection of trends in long-term correlated records, Physica A, 338, 1687–1695, 2009.

Tabeling P.: Two-dimensional turbulence: a physicist approach, Phys. Reports, 362, 1–62, 2002.

von Storch, H. and Zwiers, F.: Statistical Analysis in Climate Research, Cambridge University Press, 1999.

White, H. and Granger, C. W. J.: Consideration of trends in time series, J. Time Series Econometrics, 3(1), Article 2, doi:10.2202/1941-1928.1092, 2011.

Wilcox, R. R.: Applying Contemporary Statistical Techniques, Academic Press, 2003.

Wu, W. B. and Zhao, Z.: Inference of trends in time series, J. R. Statist. Soc. B, 69, 391-410, 2007.

Wyngaard, J. C. and Weil, J. C.: Transport asymmetry in skewed turbulence, Phys. Fluids A, 3, 155-162, 1991.