Extreme Value Statistics of the Total Energy in an Intermediate-Complexity Model of the Midlatitude Atmospheric Jet. Part I: Stationary Case

MARA FELICI

PASEF, Dipartimento di Matematica ed Informatica, Università di Camerino, Camerino, and Dipartimento di Matematica U. Dini, Università di Firenze, Florence, Italy

VALERIO LUCARINI, ANTONIO SPERANZA, AND RENATO VITOLO

PASEF, Dipartimento di Matematica ed Informatica, Università di Camerino, Camerino, Italy

(Manuscript received 23 January 2006, in final form 29 June 2006)

ABSTRACT

A baroclinic model of intermediate complexity for the atmospheric jet at middle latitudes is used as a stochastic generator of atmosphere-like time series. In this case, time series of the total energy of the system are considered. Statistical inference of extreme values is applied to sequences of yearly maxima extracted from the time series in the rigorous setting provided by extreme value theory. The generalized extreme value (GEV) family of distributions is used here as a basic model, both for its qualities of simplicity and its generality. Several physically plausible values of the parameter T_E, which represents the forced equator-to-pole temperature gradient and is responsible for setting the average baroclinicity in the atmospheric model, are used to generate stationary time series of the total energy. Estimates of the three GEV parameters—location, scale, and shape—are inferred by maximum likelihood methods. Standard statistical diagnostics, such as return level and quantile–quantile plots, are systematically applied to assess goodness-of-fit. The GEV parameters of location and scale are found to have a piecewise smooth, monotonically increasing dependence on T_E. The shape parameter also increases with T_E but is always negative, as is required a priori by the boundedness of the total energy. The sensitivity of the statistical inferences is studied with respect to the selection procedure of the maxima: the roles occupied by the length of the sequences of maxima and by the length of data blocks over which the maxima are computed are critically analyzed. Issues related to model sensitivity are also explored by varying the resolution of the system. The method used in this paper is put forward as a rigorous framework for the statistical analysis of extremes of observed data, to study the past and present climate and to characterize its variations.

1. Introduction

The study of climatic extreme events is of paramount importance for society, particularly in the fields of engineering as well as environmental and territorial planning. Indeed, temporal variations in the statistics of extreme events may have effects that are more acute and disruptive than changes in the mean climate (Katz and Brown 1992). In works of economical nature (see, e.g., Nordhaus 1994; Kunkel et al. 1999), the special role played by the extreme events in terms of impacts is modeled by the hypothesis that the costs associated with climatic change can be represented as strong nonlinear functions of the observed variations in surface temperature. This constitutes a clear motivation why, when the impacts of climatic change are examined, the interest for variations in the statistics of extreme events plays a strategic role (Watson et al. 2001; Lucarini 2002). Apart from the evaluation of costs due to wind storms (Rootzén and Tajvidi 2001), more recently estimates of wind speed extremes have proved relevant in the evaluation of potential production of wind energy in a considered region (Mortensen et al. 1993; Lavagnini et al. 2006).

Recently, Karl and coauthors (Karl et al. 1996; Karl

Corresponding author address: Dr. Renato Vitolo, Department of Mathematics and Informatics, University of Camerino, via Madonna delle Carceri, 62032 Camerino (MC), Italy.
E-mail: renato.vitolo@unicam.it

DOI: 10.1175/JAS3895.1

and Knight 1998) analyzed in qualitative terms the existence of trends in the frequency of extreme (precipitation) events. Here the authors stated that "the percentage of the United States with a much above normal proportion of total annual precipitation from extreme precipitation events (daily events at or above 2 in.)" showed an increase from 9% in 1910–20 to about 11% in the 1990s. Despite scientific criticism of these papers by other researchers in the field, the basic idea that the frequency of extreme events may change together with average surface temperature has been increasingly discussed, eventually becoming one of the issues debated by the Intergovernmental Panel on Climate Change (IPCC): a specific report on *changes* in extreme weather and climate events was issued in 2002 (available online at http://www.ipcc.ch/pub/support.htm). When dealing with extremes of complex processes, basic questions to be asked are what is the correct way of measuring extremes? Are we concentrating on *local* or *global* fluctuations of the system in question? How do we measure local extremes? Extremes of wind speeds, rainfall amounts, or economical damage? Moreover, the enhancement in the extreme events might be quantified either in terms of the number of events or in the size of the average extreme event or a combination thereof. Several other ambiguities often make the literature on the subject difficult to follow.

Two important weaknesses of much of the work on the subject of extreme meteoclimatic events and their trends are

- the lack of interpretation of the dynamical mechanisms that are supposed to cause the hypothesized changes in the probability distribution of extremes; these mechanisms are often just alluded to, instead of being explicitly formulated quantitatively and analyzed.
- the lack of a common and theoretically founded definition of extremes.

The deficit mentioned in the first point may have a negative effect on both deterministic and statistical studies of the phenomena in question. One major example concerning global processes is that, despite the great attention given to the subject, very few researchers have investigated the basic mechanisms that should associate an increased CO_2 concentration to enhanced extreme weather events in detail, especially in the case of extratropical cyclones (see, e.g., Lionello et al. 2002). Regarding the observed climate change, as summarized from chapter 2 of the 2001 Working Group I report of the IPCC (Houghton et al. 2001), while several studies at regional level claim that increases in extratropical cyclones seem to have occurred in several regions of the Northern Hemisphere (but *not* in the Southern Hemisphere!), the mechanisms involved are *not* clear, and it is not certain whether these trends are multidecadal fluctuations or rather part of a longer-term trend. In fact, we may guess that in the context of the current and future global warming trends (if taken for granted) two contrasting mechanisms might be at work: on the one hand, the fact that the atmosphere and surface waters are warming up might allow for a moistening of the atmosphere, thus allowing for increases in stored energy in the form of latent heat; but at the same time, a moderating influence might be exerted by the polar warming, which reduces the equator-to-pole temperature difference, thus decreasing the average baroclinicity of the system. Some parts of this complex dynamical *chain* have been analyzed: for example, Allen and Ingram (2002) postulate that the changes in precipitation extremes are controlled by the Clausius–Clapeyron relationship. Zwiers and Kharin (1998) found a decrease in extreme wind speed under CO_2 doubling in a GCM simulation. Nevertheless, to our knowledge, the complete picture is far from being understood.

As for the second point above, the lack of a common, rigorous framework for the statistical analysis of extremes (with exceptions such as, e.g., Katz et al. 2002; Zwiers and Kharin 1998; Kharin and Zwiers 2000) provides a serious drawback for the interpretation and comparison of results from different studies. This problem is not justified since mathematical theories on extreme events are well developed (Castillo 1988; Coles 2001; Embrechts et al. 1997; Fisher and Tippett 1928; Galambos 1978; Gnedenko 1943; Lindgren et al. 1983) and the derived methods are quite successful in many applications (Katz et al. 2002; Perrin et al. 2006; Zwiers and Kharin 1998; Kharin and Zwiers 2000). One basic ingredient of the theory is Gnedenko's theorem (Gnedenko 1943), which states that, under fairly mild assumptions, the distribution of the block maxima of a sample of independent identically distributed variables converges to a member of a parametric family of distributions: the so-called generalized extreme value (GEV) family. Note that one of the earliest applications of this theory in the natural sciences occurred specifically in a meteoclimatic setting (Jenkinson 1955). Other statistical models for extreme events include the *r*-largest statistics, threshold exceedance models, such as the generalized Pareto distribution, and point processes (see Coles 2001).

The reliability of parametric estimates for extreme value models is highly dependant on the asymptotic nature of extreme value theory. In particular, at least

the following issues should be checked or addressed (Coles 2001):

1) *Independence* of the selected extreme values
2) Using a *sufficiently large* number of extremes
3) Using values that are *genuinely* extreme

Despite the importance of the third requirement, many studies actually deal with so-called soft extremes (Klein Tank and Können 2003), which are maxima of blocks that are either too short or values having return periods that are too small to allow the basic assumptions of the theory to hold. This is often the consequence of the limited amount of available data: on the one hand, one has to restrict to maxima of data blocks, thereby discarding *most* available data; on the other hand, one would like to have a *long* sequence of extreme values. The net result is that the assumptions of the extreme value theorems often go unchecked and are, at times, plainly impossible to check, since the available systematic climatic records cover the last century at best. Therefore, thinking in terms of annual maxima, in such cases we only have 100 extremes. Adapting the definition of extremes to the needs of the work means that the reliability of the resulting estimates is seriously reduced.

The goal of this paper is to infer and critically check the statistics of extreme values, in the GEV framework, on the atmosphere-like time series produced by a dynamical system describing the midlatitude atmospheric circulation. This system displays internally generated noise (a chaotic attractor) and is used as a *stochastic generator* of data. We consider the time series of the system total energy $E(t)$, which is a relevant physical quantity of global character. Subsequently, we analyze how the GEV distribution inferred from block maxima of $E(t)$ depends on the value of the forced equator-to-pole temperature difference T_E, which controls the baroclinicity of the model. The reliability of the GEV fits is examined by considering both shorter sequences of extremes and so-called soft extremes. Moreover, issues related to model error and sensitivity are briefly examined by analyzing the effects of variations in model resolution. The usage of numerically generated data allows us to avoid all the difficulties associated with the available climatic records, such as missing observations and low-quality data. In particular, we do not need to worry about the wastage of data caused by the selection of annual maxima, which is a serious limitation when considering observed data. In the methodological sense, and as far as statistical inference is concerned, our approach is similar to that of Zhang et al. (2004). However, an important difference is that the statistics of the time series $E(t)$ generated by the atmospheric model *cannot be directly chosen:* there is no explicit formula relating the probability density function of $E(t)$ and the parameter T_E.

The structure of the paper is thus outlined. In section 2 we first describe the setup of the numerical experiments performed with the atmospheric model and then the adopted methods of statistical analysis of extreme values. The results for the considered reference case of 1000 yearly maxima are presented in section 3. Sensitivity of the inferences is assessed in section 4 by varying the length of yearly maxima sequences, the block length over which maxima are taken, and the model resolution. The dependence of the GEV parameters with respect to T_E is also analyzed in this section. Section 5 summarizes the results and their relation to the above discussion. The model of the baroclinic jet used as a stochastic generator is described in the appendix, referring to Lucarini et al. (2005) for a thorough discussion.

2. Data and methods

a. Total energy of the atmospheric model

We consider a quasigeostrophic intermediate complexity model (Speranza and Malguzzi 1988; Malguzzi et al. 1990; Lucarini et al. 2005; also see the appendix) providing a basic representation of the turbulent jet and of the baroclinic conversion, barotropic stabilization, thermal diffusion, and viscouslike dissipation processes, which characterize the physics of the atmospheric circulation in the midlatitudes. The model is relaxed toward a given equator-to-pole temperature profile, which acts as baroclinic forcing. It features several degrees of freedom in the latitudinal direction and two layers in the vertical, the minimum for baroclinic conversion to take place (Pedlosky 1987; Phillips 1954). The system's statistical properties change quite relevantly when the parameter T_E, determining the forced equator-to-pole temperature gradient, is varied. In particular, as T_E increases, we go from a stationary to an atmosphere-like chaotic regime with internally generated noise. By chaotic, we mean that the system possesses a strange attractor in phase space (Eckmann and Ruelle 1985). For a detailed description of the model physics and dynamics see Lucarini et al. (2005).

In the present setting, the model is used as a stochastic generator of time series of the total energy, both for testing the reliability of different statistical approaches (cf. with Zhang et al. 2004) and for studying the dependence of extremes from the parameter T_E. A uniformly spaced grid of 21 values of T_E is fixed in the range [10, 50], starting from 10 and increasing with step 2. The baroclinic model is run for T_E fixed at each of these

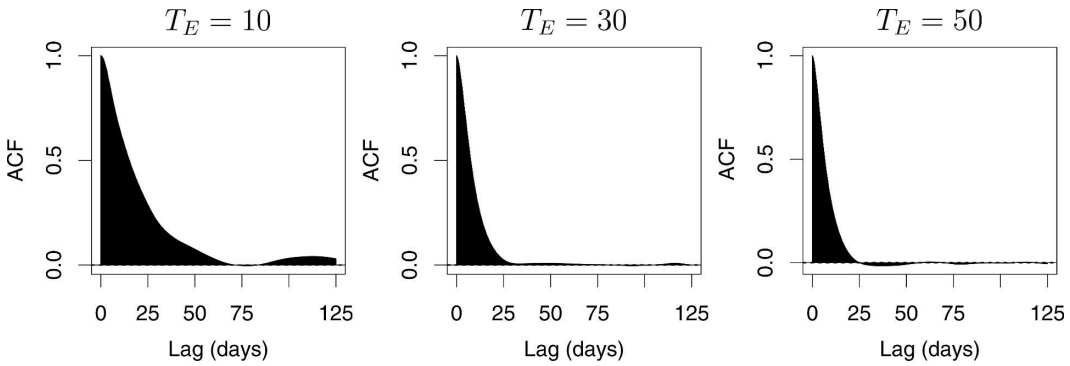

FIG. 1. Autocorrelations of the total energy time series for (left) $T_E = 10$, (center) $T_E = 30$, and (right) $T_E = 50$; time lag in days is on the horizontal axis. The full 6-hourly time series of 1000 yr have been used (see section 2a).

values, producing 21 simulations, which are 1000 yr of length (preceded by an initial, discarded transient of 5 yr) where the total energy $E(t)$ is recorded every 6 h [the formula of the total energy is given in the appendix, Eq. (31)]. We recall that, in the nondimensionalization of the system, $T_E = 1$ corresponds to 3.5 K, 1 unit of total energy corresponds to roughly 5×10^{17} J, and $t = 0.864$ is 1 day; see Lucarini et al. (2005) for details.

For each of the selected values of T_E, a chaotic attractor is numerically detected in the phase space of the model. This is illustrated by the autocorrelations of the time series of the total energy $E(t)$ (Fig. 1), which decay to zero on a time scale that is comparable with that of the atmospheric system [roughly 10–15 days (Lorenz 1967)]. Since all parameters of the model are kept fixed in each simulation, upon discarding the initial transient the time series of $E(t)$ may be considered a realization of a stationary stochastic process.

The distribution of the total energy time series is visualized by means of the histograms and boxplots in Fig. 2, for three values of T_E. Notice that, as T_E increases,

- the upper tail of the distribution becomes heavier, whereas the lower tail shortens; and
- both the average value and the variability of the total energy time series increase.

The latter point is clearly visualized in Fig. 3, where the time-averaged total energy is displayed for each of the 21 stationary time series, together with confidence intervals. Throughout the paper, confidence intervals are computed as average plus/minus sample standard deviations multiplied by 1.96.

In concluding this section a theoretical remark is in order. All the strange attractors examined are implicitly *assumed* to possess a unique Sinai–Ruelle–Bowen (SRB) ergodic invariant measure (Eckmann and Ruelle 1985). This is indeed a rather general and difficult problem in dynamical systems and physics. On the one hand, existence of a unique SRB measure is *necessary* to correctly associate a stationary stochastic process with the dynamical evolution law. On the other hand, existence of a unique SRB measure is a very strong regularity assumption for a dynamical system: it is not even known whether invariant measures exist at all and, if so,

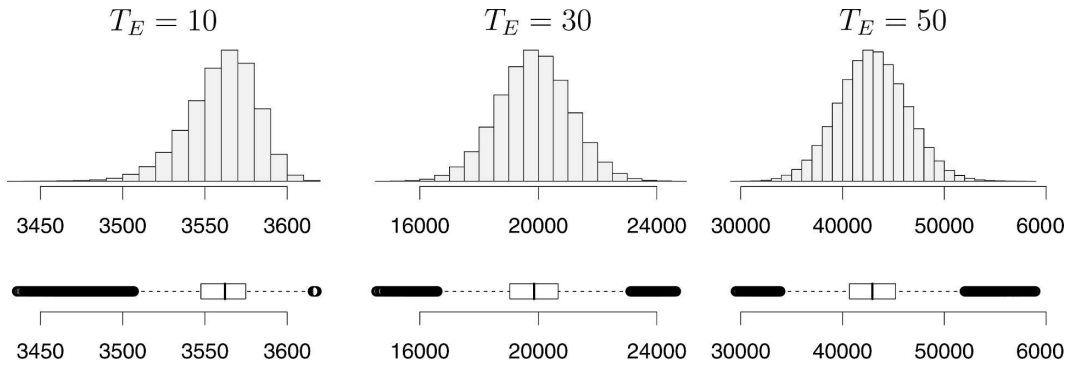

FIG. 2. Histograms and boxplots of the total energy time series for (left) $T_E = 10$, (center) $T_E = 30$, and (right) $T_E = 50$.

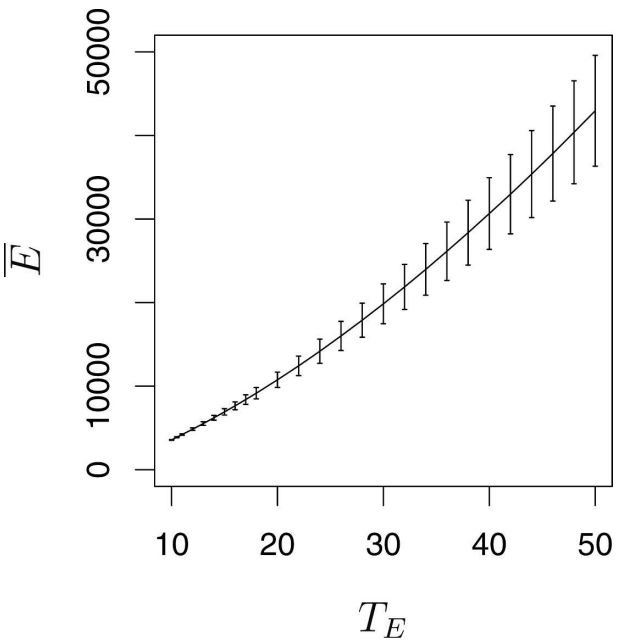

FIG. 3. Time-averaged total energy $\overline{E}$ (vertical axis) as a function of T_E (horizontal axis) for each of the 21 selected values of T_E. Confidence bands (average ± 1.96 times sample std dev) are added. The full 6-hourly time series of 1000 yr have been used (see section 2a).

whether a finite or infinite number of invariant measures coexist for a given chaotic system. Moreover, even if an SRB measure exists and is unique, it is in general nonparametric: there is *no* explicit formula relating the statistical behavior to the system's equation and parameters. Our assumption of existence and uniqueness of a SRB measure is coherent with the chaotic hypothesis proposed by Gallavotti and Cohen (Gallavotti and Cohen 1995a,b; Gallavotti 1996; Cohen and Gallavotti 1999).

b. *Parameter estimation and model assessment in GEV inference*

As discussed in the previous section, the time series we work with are characterized by fast decay of autocorrelations (roughly 10–15 days), which implies weak (short time range) dependence of the observations (cf. Fig. 1). Inference of threshold exceedance models (Coles 2001; Embrechts et al. 1997; Lindgren et al. 1983) is, in this case, complicated by the choices of suitable threshold values and cluster size for declustering (see, e.g., Coles 2001, chapter 5), which might be somewhat arbitrary in the applications. On the other hand, since the dependence is short range, if the maxima of the total energy time series are taken over sufficiently large data blocks, then they may be considered independent with good approximation. This is why we have preferred the GEV to threshold models. Moreover, since we can generate time series of arbitrary length, for simplicity we refrained from using the *r*-largest statistics, which are often a valid alternative to the GEV, especially when data scarcity is an issue. In this section, therefore, we recall the methods of GEV inference as far as needed in the present work. The exposition is largely based on Coles (2001). Also see Castillo (1988); Coles (2001); Embrechts et al. (1997); Fisher and Tippett (1928); Galambos (1978); Gnedenko (1943); and Lindgren et al. (1983) for methodology and terminology of extreme value theory.

Gnedenko's theorem (Gnedenko 1943), or the three types theorem, first presented in a slightly less general form by Fisher and Tippett (1928), states that, under fairly mild assumptions, the distribution of the block maxima of a sample of independent identically distributed variables converges, in a suitable limit, to one of three types of extreme value distributions. The three types are in fact special cases of the GEV family of distributions (also called von Mises type):

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \qquad (1)$$

for x in the set $\{x: 1 + \xi(x - \mu)/\sigma > 0\}$ and $G(x) = 0$ otherwise, with $-\infty < \mu < +\infty$, $\sigma > 0$, and $-\infty < \xi < +\infty$. The quantities (μ, σ, ξ) are called location, scale, and shape parameter, respectively. In such a framework, statistical inference of extreme values amounts to estimating the GEV distributional parameters (μ, σ, ξ) for a given time series and assessing the quality of the fit. If $\xi > 0$ ($\xi < 0$) the distribution is usually referred to as Fréchet (Weibull) distribution, whereas if $\xi = 0$ we are experiencing the Gumbel distribution. See Embrechts et al. (1997); Castillo (1988); Coles (2001); Galambos (1978); and Lindgren et al. (1983) for details and examples.

In practical applications of the extreme value theory the distribution function of the data (the *parent* distribution) typically is unknown. Therefore, both the type of limiting distribution and the parameter values must be inferred from the available data and the quality of the resulting estimates should always be assessed. For GEV inference, a sequence of maxima is constructed by subdividing the available data $\{x_i\}$ into blocks of equal length and by extracting the maximum from each block. The block length is one of the choices that plays the usual, critical role between bias and variance in the parametric estimates. On the one hand, by using shorter blocks a longer sequence of maxima is obtained, resulting in smaller uncertainties for the infer-

ences. At the same time, approximation to the limiting distribution might be worse due to the introduction of a bias. On the other hand, if the blocks are too long an enhanced uncertainty is induced for the inferred values of the GEV parameters. In many situations concerning climate studies a reasonable (and sometimes compulsory) choice is to consider the annual maxima (see Coles 2001).

Assume that the observations in the time series are equispaced in time and that none of them is missing [both conditions are often violated in concrete cases, see, e.g., Perrin et al. (2006)]. Let n be the number of observations in a year and denote $M_{n,1}, \ldots, M_{n,m}$ as the sequence of the annual maxima, that is, the maxima over the consecutive data blocks of length n. If the variables X_n are independent, then the variables $M_{n,1}, \ldots, M_{n,m}$ are independent as well. In fact, approximate independence of the $M_{n,i}$ holds also in the case of weak-dependent stationary sequences, see Lindgren et al. (1983) and Coles (2001) for definitions and examples.

Among the numerous methods to infer the GEV parameters [moment based or graphical techniques, see Castillo (1988)], we adopt the maximum likelihood estimator for its great adaptability to changes of models. Denote $\boldsymbol{\theta} = (\mu, \sigma, \xi)$ as the parameter vector for the GEV density $g(x; \boldsymbol{\theta})$, the latter being the derivative of $G(x) = G(x; \boldsymbol{\theta})$ in Eq. (1). In the stationary context, the block maxima of the observed data are assumed to be realizations of a stationary stochastic process with density $g(x; \boldsymbol{\theta}^0)$, where $\boldsymbol{\theta}^0$ is the unknown parameter vector. The *maximum likelihood estimator* $\hat{\theta}^0$ of $\boldsymbol{\theta}^0$ is defined as the value that maximizes the likelihood function

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} g(M_{n,i}; \boldsymbol{\theta}). \tag{2}$$

To put it simply, maximizing $L(\boldsymbol{\theta})$ yields the parameter values for which the probability of observing the available data is the highest. It is often more advantageous to maximize the *log-likelihood function*

$$l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta}) = \sum_{i=1}^{m} \log g(M_{n,i}; \boldsymbol{\theta}), \tag{3}$$

and, according to Eq. (1), the log-likelihood function $l(\mu, \sigma, \xi)$ is given by

$$-m \log\sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} \left\{ \log\left[1 + \xi\left(\frac{M_{n,i} - \mu}{\sigma}\right)\right] - \left[1 + \xi\left(\frac{M_{n,i} - \mu}{\sigma}\right)\right]^{-(1/\xi)} \right\} \tag{4}$$

if $\xi \neq 0$ and by

$$-m \log\sigma - \sum_{i=1}^{m} \left\{ \left(\frac{M_{n,i} - \mu}{\sigma}\right) - \exp\left[-\left(\frac{M_{n,i} - \mu}{\sigma}\right)\right]\right\} \tag{5}$$

if $\xi = 0$, defined on the points $M_{n,i}$ that, in the case $\xi \neq 0$, satisfy the condition $1 + \xi(M_{n,i} - \mu)/\sigma > 0$ for all $i = 1, \ldots, m$.

Approximate confidence intervals for $\hat{\boldsymbol{\theta}}^0$ are constructed using the fact that each component of $\hat{\boldsymbol{\theta}}^0 = (\hat{\theta}_1^0, \hat{\theta}_2^0, \hat{\theta}_3^0) = (\hat{\mu}^0, \hat{\sigma}^0, \hat{\xi}^0)$ is asymptotically normal (Coles 2001):

$$\hat{\theta}_i^0 \sim N(\theta_i^0, \hat{\psi}_{i,i}) \quad \forall i = 1, \ldots, d, \tag{6}$$

where $N(a, b)$ denotes the normal distribution with mean a and variance b and $\hat{\psi}_{i,j}$ is a generic element of the inverse of the observed information matrix $\mathbf{l}_0(\boldsymbol{\theta})$ defined by

$$\mathbf{l}_0(\boldsymbol{\theta}) = \left(-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right)_{i,j} \quad \forall i, j = 1, \ldots, d \tag{7}$$

and evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^0$. From Eq. (6) one obtains the $(1 - \alpha)$ confidence interval for $\hat{\theta}_i^0$:

$$\theta_i^0 \pm z_{\alpha/2} \sqrt{\hat{\psi}_{i,i}}, \tag{8}$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution. All confidence intervals in this paper are computed by Eq. (8), except when a more detailed analysis is presented. For example, in the assessment of inference quality, confidence intervals are also computed by a standard bootstrap procedure (applied to the sequence of annual maxima) and by profile likelihood. The latter technique consists of the following. Consider the parameter ξ, for example. The profile likelihood of ξ is obtained by setting μ and σ to their maximum likelihood estimates, $\hat{\mu}^0$ and $\hat{\sigma}^0$, respectively, in the log-likelihood function l [Eq. (4)]. The profile likelihood plot is the graph $(x, y) = (\xi, l(\hat{\mu}^0, \hat{\mu}^0, \xi))$, giving a section of the likelihood surface as viewed from the ξ axis. A confidence interval can be computed by determining the intersections of the horizontal line

$$y = \widehat{\xi^0} - \frac{1}{2} q_{0.95} \tag{9}$$

with the profile likelihood graph, where $q_{0.95}$ is the 95% quantile of the χ^2 distribution with 1 degree of freedom. See Coles (2001, section 2.6.6 and 2.7) for theory and examples.

One of the main goals of extreme value theory is

estimating the probability of occurrence of events that are *more* extreme than those that have been observed thus far. Let z_p be the value that has a probability p to be exceeded every year by the annual maximum $\Pr\{M_{n,i} > z_p\} = p$ with $0 < p < 1$. In common terminology z_p is called the *return level* associated with the *return period* $1/p$. A maximum likelihood estimator for z_p is obtained by plugging the estimates for $\hat{\boldsymbol{\theta}}^0 = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ into the quantiles of $G(x)$, obtained by inverting Eq. (1). This yields the estimator

$$\hat{z}_p = \begin{cases} \hat{\mu} - \dfrac{\hat{\sigma}}{\hat{\xi}}\{1 - [-\log(1-p)]^{-\hat{\xi}}\} & \text{for} \quad \hat{\xi} \neq 0 \\ \hat{\mu} - \hat{\sigma}\log[-\log(1-p)] & \text{for} \quad \hat{\xi} = 0. \end{cases} \tag{10}$$

The variance of the return level estimator $\hat{z}_p$ is approximated as

$$\mathrm{Var}(\hat{z}_p) \approx \boldsymbol{\nabla} z_p^{\mathrm{T}} \mathbf{V} \boldsymbol{\nabla} z_p, \quad \text{where}$$
$$\boldsymbol{\nabla} z_p^{\mathrm{T}} = \left(\frac{\partial z_p}{\partial \mu}, \frac{\partial z_p}{\partial \sigma}, \frac{\partial z_p}{\partial \xi} \right), \tag{11}$$

$\mathbf{V}$ is the variance–covariance matrix

$$\mathbf{V}(\mu, \sigma, \xi) = \begin{pmatrix} \mathrm{Var}(\mu) & \mathrm{Cov}(\mu, \sigma) & \mathrm{Cov}(\mu, \xi) \\ \mathrm{Cov}(\sigma, \mu) & \mathrm{Var}(\sigma) & \mathrm{Cov}(\xi, \sigma) \\ \mathrm{Cov}(\xi, \mu) & \mathrm{Cov}(\xi, \sigma) & \mathrm{Var}(\xi) \end{pmatrix}, \tag{12}$$

and both $\boldsymbol{\nabla} z_p$ and $\mathbf{V}$ are evaluated at the maximum likelihood estimate $\hat{\boldsymbol{\theta}}^0 = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$. This allows for the construction of confidence intervals for $\hat{z}_p$ and is referred to as the delta method. Again, profile likelihood and a bootstrap technique are used for goodness-of-fit assessment of the return level inferences.

Only Weibull distributions ($\xi < 0$) allow for the possibility of having $p = 0$, corresponding to a return level with an infinite return period. In this case,

$$z_0 = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}. \tag{13}$$

Information on the return levels is usually reported in the *return level plot*, where $\hat{z}_P$ is plotted against $\log y_p$, where $y_p = -\log(1 - p)$ [cf. Eq. (10)]. The return level plot is linear for the Gumbel distribution, concave for $\xi > 0$ (Frechet), and has the horizontal asymptote Eq. (13) for $\xi < 0$ (Weibull). Note that the smallest values of p are usually those of interest, since they correspond to very rare (*particularly* extreme) events. In the return level plots, events with a short return period (large probability p) are compressed near the origin of the axes, while outliers and rare events (small p) are highlighted. It is for this reason that such plots are very useful tools for both model analysis and diagnosis.

The above procedures for inference require assessment with reference to the available data. Useful graphical checks are the probability plot, the quantile–quantile plot (QQ plot), and the return level plot. The first is the comparison between the estimated and the empirical distribution function $\tilde{G}(x)$: the latter is a step-function defined by

$$\tilde{G}(M_{(i)}) = \frac{i}{m+1}, \tag{14}$$

where $M_{(i)}$ is the order statistics for the sequence $M_{n,1}, \ldots, M_{n,m}$ of m block maxima. Note alternative definitions of the empirical distribution function exist, see Castillo (1988).

The QQ plot, formed by the points

$$\left\{ \left(\tilde{G}^{-1}\left(\frac{i}{m+1} \right), m_{(i)} \right), \quad \forall i = 1, \ldots, m \right\}, \tag{15}$$

highlights the behavior of the model tail, which is often the most interesting part. Substantial departures of the above-listed plots from the diagonal indicate inadequacy of the GEV model or other systematic errors. Another diagnostic plot is constructed by adding confidence intervals for $\hat{z}_p$ and return levels of the empirical distribution function, according to Eq. (15), to the return level plot (see above). Agreement of the empirical distribution function with the return level curve suggests goodness-of-fit and adequacy of the GEV model.

All computations and plots in this paper have been carried out with the statistical software R (Ihaka and Gentleman 1996) freely available at www.r-project.org under the General Public License (GPL). The library ismev (www.cran.r-project.org), which is an R-port of the routines written by Stuart Coles as a complement to Coles (2001), has been used with minor modifications.

3. GEV inferences for 1000 annual maxima

The annual maxima are extracted from the 6-hourly time series of the energy described in section 2. Each series contains $4 \times 365 \times 1000 = 1\,460\,000$ data. We fix $n = 1460$ as the length of the data blocks over which the maxima $\{M_{n,i} | i = 1, \ldots, 1000\}$ are computed. Thereby, for each value of T_E we obtain a sequence of 1000 annual extremes of the total energy. The yearly maxima are linearly uncorrelated (Fig. 4), suggesting that it is both safe and reasonable to assume weak dependence. This can also be compared with the autocorrelation decay time in Fig. 1.

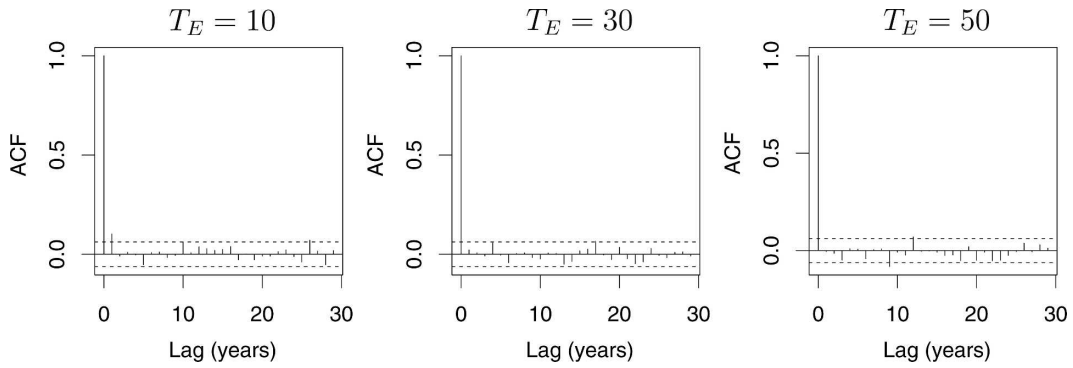

FIG. 4. Autocorrelations of the sequences of 1000-yearly maxima of the total energy time series for (left) $T_E = 10$, (center) $T_E = 30$, and (right) $T_E = 50$.

On theoretical grounds we can deduce one constraint on the distribution of extremes for the total energy time series. Indeed, since the attractor is contained within a bounded domain of the phase space and since the energy observable $E(t)$ defined in Eq. (A15) is a continuous function of the phase space variables, it turns out that the total energy is bounded on any orbit belonging to (or converging on) the attractor. Therefore, the extremes of the total energy are *necessarily* Weibull distributed (ξ is negative). This provides a theoretically founded criterion for quality assessment of the GEV inferences. We note that such a strict Weibull constraint is specific to the present setup, where a global observable (the total energy) obeying global balances is used. Although we might expect that the total energy extremes of any atmospheric model should be Weibull distributed, for other meteoclimatic variables this might be not the case. For example, wind speeds in extratropical latitudes are known to be approximately Weibull distributed, although a Gumbel fit ($\xi = 0$) often performs better on extreme wind speeds (Perrin et al. 2006). Hydrological variables, such as precipitation (Smith 2006) and streamflow (Morrison and Smith 2002), often display heavy tails ($\xi > 0$). See Katz et al. (2002) and references therein.

The GEV parameters (μ, σ, ξ) are estimated by the maximum likelihood method (see section 2b) from the sequences of yearly maxima. The fitted values of (μ, σ, ξ), together with confidence bands [computed by the observed information matrix, Eq. (8)], are plotted as functions of T_E in Fig. 5. The inferred parameters μ and σ increase monotonically with T_E. Estimates of ξ are negative in each case and the related confidence intervals are markedly bounded away from zero: observed information matrix, profile likelihood, and bootstrap yield similar estimates. The theoretical expectation of Weibull distribution is thus confirmed. Also notice that the uncertainty in ξ may reach up to 21% of its value, whereas the parameters μ and σ are quite accurately estimated: the maximal uncertainties in μ and σ are 0.1% and 2.5% of the corresponding value, respectively.

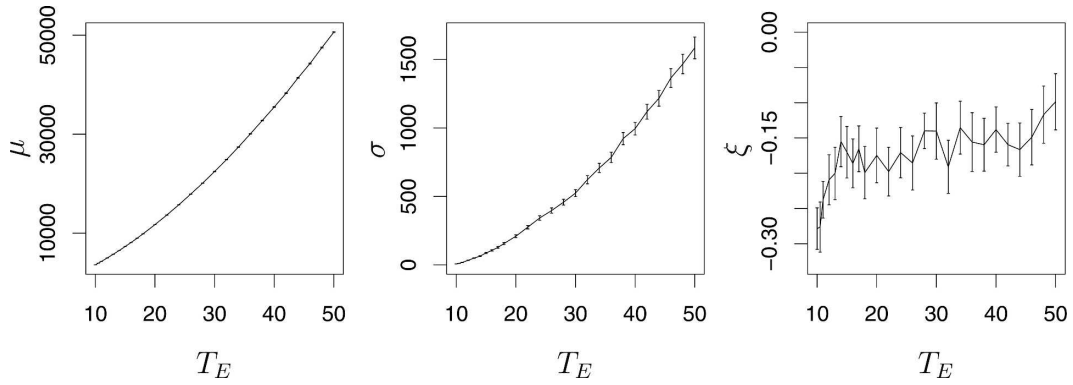

FIG. 5. Maximum likelihood estimates of (left) μ, (center) σ, and (right) ξ (on vertical axis) for each of the 21 sequences of 1000 maxima of the total energy, against the corresponding values of T_E (horizontal axis). Confidence intervals are added with error bars but are hardly visible for (left) μ at the selected scale.

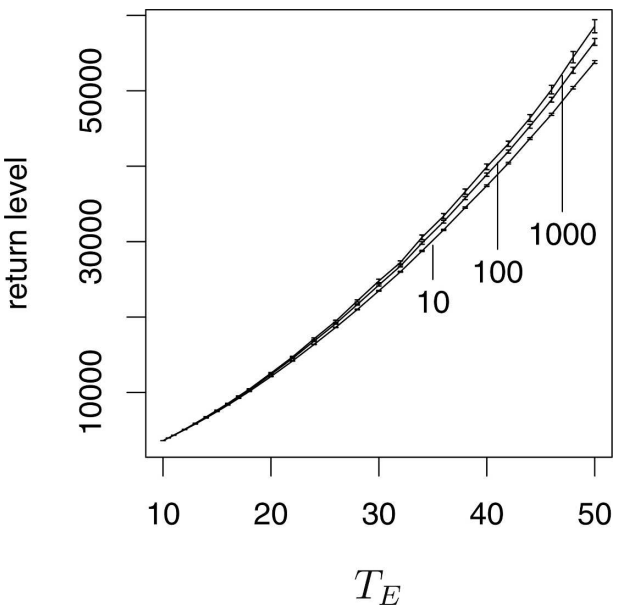

FIG. 6. Maximum likelihood estimates of the 10-, 100-, and 1000-yr return levels of the total energy for each of the 21 stationary series of 1000 maxima of the total energy, against the corresponding values of T_E (horizontal axis). Confidence intervals are added with error bars but are hardly visible (at the selected scale).

Information on the tails of the energy distribution is expressed in a straightforward manner by the return level plots (see previous section for the definition). In Fig. 6, return levels with return periods of 10, 100, and 1000 yr are plotted as functions of T_E. Each graph is monotonically increasing with T_E and, for T_E fixed, the return levels increase with the return period.

The dependence of the GEV probability density with respect to T_E is illustrated in Fig. 7. The increase of scale and location parameters with T_E induces a rightward shift and a broadening of the probability density function. In particular, from the geophysical point of view, the range of possible extreme values of the total energy expands with T_E and the magnitude increases too. In fact, this behavior sets in for T_E right after the creation of the chaotic attractor (see Fig. 7, right panel).

Smoothness of GEV inferences with respect to system parameters

The dependence on T_E of the time-averaged total energy and the inferred GEV parameters (including the return levels) is rather smooth (see Figs. 3 and 5). This strongly suggests the existence of functional relations of the form

$$\mu = \alpha_\mu T_E^{\gamma_\mu} \quad \text{and} \quad \sigma = \alpha_\sigma T_E^{\gamma_\sigma}. \tag{16}$$

Such power laws are fitted to the graphs of μ and σ as follows.

To illustrate, consider μ and denote $\hat{\mu}(T_E^j)$ and $s_{\hat{\mu}}(T_E^j)$ as the maximum likelihood estimate of μ and the related standard deviation (calculated by the observed information matrix), respectively, where T_E^j is one of the 21 chosen values in the interval [10, 50]. A bootstrap procedure is performed where iterated realizations of a sequence of 21 independent Gaussian variables with mean $\hat{\mu}(T_E^j)$ and standard deviation $s_{\hat{\mu}}(T_E^j)$ are simulated. For each realization, a power-law fit as in Eq. (16) is performed. The sample average and standard deviation of the obtained fits, constructed independently for μ and σ, are reported in Tables 1 and 2. Two distinct ranges of T_E are identified, where μ scales by a different exponent ($\gamma_{\mu,1}$ and $\gamma_{\mu,2}$, respectively) (see also Fig. 8, left panel). For $T_E \lesssim 18$, $\gamma_{\mu,1} \sim 1.73$ while $\gamma_{\mu,2} \sim 1.6$ for $T_E \gtrsim 18$. The time-mean total energy of

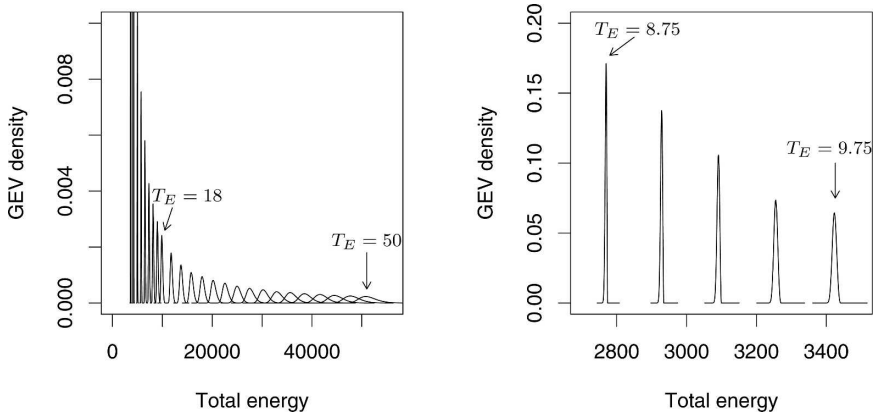

FIG. 7. (left) Probability density functions of the GEV for the 21 values of T_E in the considered range [10, 50] and for the additional values $T_E = 8.75$, 9, 9.25, 9.5, and 9.75. (right) Same as left, but for $T_E = 8.75$, 9, 9.25, 9.5, and 9.75.

TABLE 1. Power-law fits of the location parameter μ as a function of T_E of the form $\mu \propto T_E^{\gamma}$, performed in two adjacent intervals of T_E. The number of used annual extremes is n, and T_E^b is the value of T_E separating the two intervals. Compare with Fig. 8.

n	$\gamma_{\mu,1}$	T_E^b	$\gamma_{\mu,2}$
1000	1.7310 ± 0.0007	18	1.6019 ± 0.0017
300	1.7311 ± 0.0013	18	1.6017 ± 0.0030
100	1.7219 ± 0.0018	17	1.5982 ± 0.0018

TABLE 2. Same as in Table 1 but for the scale parameter σ. Here the fits with exponents $\gamma_{\sigma,1}$ and $\gamma_{\sigma,2}$ hold for T_E such that $T_E^{b1} \leq T_E \leq T_E^{b2}$ and $T_E^{b2} \leq T_E \leq 50$, respectively. No power-law fit is found for $T_E < T_E^{b1}$. Compare with Fig. 9.

n	T_E^{b1}	$\gamma_{\sigma,1}$	T_E^{b2}	$\gamma_{\sigma,2}$
1000	15	3.011 ± 0.076	22	2.140 ± 0.025
300	14	3.236 ± 0.115	22	2.114 ± 0.047
100	14	2.944 ± 0.157	24	2.040 ± 0.093

the system has a rather similar power-law dependence on T_E (Lucarini et al. 2005). In the upper T_E range the exponent of the power law of the extremes is larger than that of the time-mean total energy (~1.52), which implies that asymptotically the extremes tend to become relatively more extreme. When considering σ, there is an initial interval of T_E where no power law is obeyed (see Fig. 9, left panel). For $22 \gtrsim T_E \gtrsim 15$, we have $\gamma_{\sigma,1} \sim 3.0$, while $\gamma_{\sigma,2} \sim 2.1$ for $T_E \gtrsim 22$. Since $\gamma_{\sigma,2} > \gamma_{\mu,2}$, for high values of T_E the broadening of distribution of the maxima tends to become consistent with respect to their average location, thus suggesting a larger variability in the maxima. Shorter sequences of yearly maxima, of length 300 and 100, lead to nearly identical estimates for both $\gamma_{\mu,j}$ and $\gamma_{\sigma,j}$, $j = 1, 2$ and for their confidence intervals, thus implying that this is a rather robust property of the system.

It turns out that analogous power-law dependence with respect to T_E is detected in the considered model for several dynamical and physical observables, such as Lyapunov dimension, maximal Lyapunov exponent, and average zonal wind (Lucarini et al. 2005). This suggests that the whole attractor of the model (more precisely, its SRB measure) obeys some scaling laws with respect to T_E. The qualitative features described above when T_E is sufficiently large, such as the form of (μ, σ) as functions of T_E and the fact that ξ seems to approach a constant negative value, are most probably related to this scaling behavior. An important question we address elsewhere is whether this is a peculiarity of the baroclinic model used here or if analogous smoothness properties are common (*generic* or *robust* in some way) for models of atmospheric dynamics, including general circulation models.

4. Sensitivity of the GEV inferences

Selecting sequences of 1000-yearly maxima results in good accuracy for the GEV inferences. The sensitivity of these estimates has been tested by relaxing the experimental conditions considered in the previous section. This has been done in several ways:

- by varying the number of extreme events (length of the sequences of yearly maxima);
- by using soft extremes (maxima are computed over data blocks corresponding to time spans shorter than 1 yr); and
- by varying the resolution of the model.

The best estimates and related uncertainties of the GEV parameters obtained under modified experimental conditions have been first compared at face value to

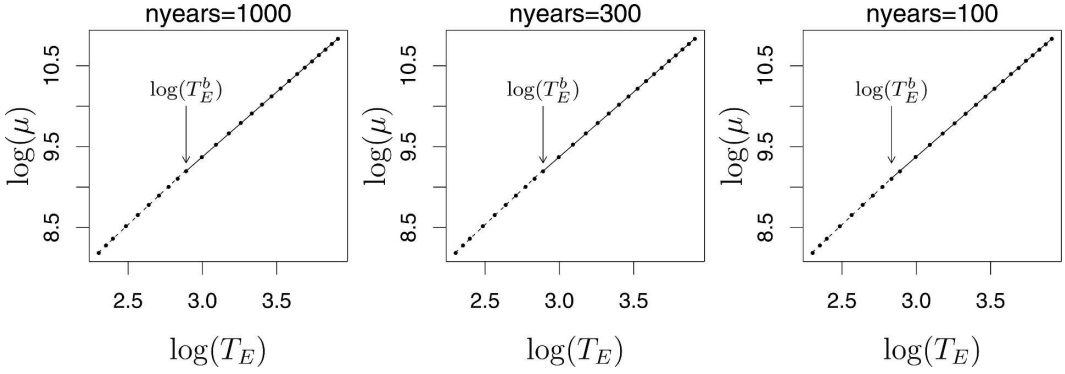

FIG. 8. Power-law fits of the inferred values of log μ (vertical axis) as a function of log T_E (horizontal axis), where (left) 1000-, (center) 300-, and (right) 100-yearly maxima have been used. In each case, there are two intervals of T_E, separated by the point T_E^b, characterized by a different scaling exponent (cf. Table 1).

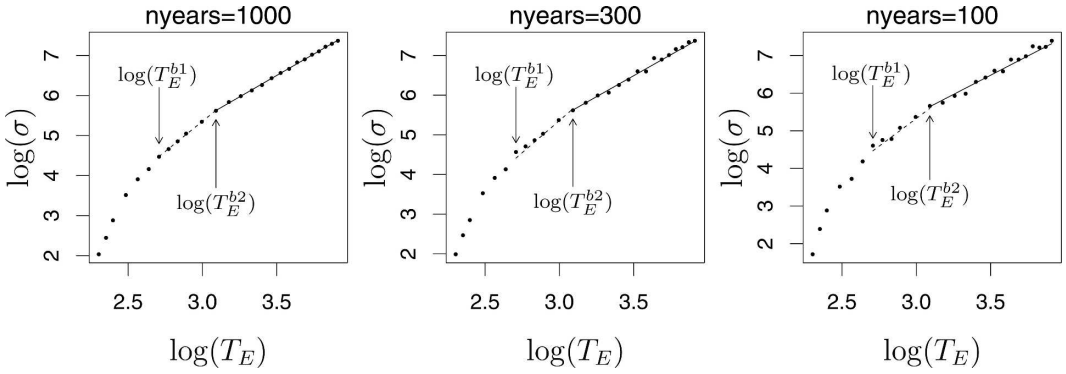

FIG. 9. Same as in Fig. 8 but for log σ. The two intervals characterized by a different scaling exponent are separated by the point T_E^{b2}, while no power law is detected for $T_E < T_E^{b1}$ (see Table 2).

that which was obtained in the reference case in order to detect mismatch due to biases and changes in precision. Moreover, the resulting differences in the GEV distributions have been inspected also by the standard graphical diagnostics, such as quantile–quantile and return level plots, and by computing bootstrap confidence intervals and profile likelihood, both for the critical parameter ξ and for the return levels.

a. Sensitivity with respect to the extreme events sample size

We now move on to the description of what is found when reducing the number of yearly maxima used for GEV inference. The, particularly unfortunate, case occurring for $T_E = 32$ is first analyzed by means of a profile likelihood plot for the GEV parameter ξ. Sequences of 1000-, 300-, 100-, and 50-yearly maxima of the total energy are used to produce the plots in Fig. 10. The cases 1000, 300, and 100 yr yield coherent estimates for ξ. For detailed diagnostics, confidence intervals are computed by the observed information matrix [Eq. (8)] and compared by those obtained by profile likelihood and by a standard bootstrap procedure. The three methods yield similar results in all cases, both for the estimates and for the confidence intervals. However, for 50 maxima the confidence intervals become very wide and a positive value for ξ is inferred, which is unphysical.

The decay of the inference quality is revealed in a different way by the profile likelihood plots for the 100-yr return levels (Fig. 11). In general it is not safe to infer, from a series of n annual extremes, return levels with return periods larger than n years. Extrapolation to larger return periods may produce incorrect values and is likely to yield significant uncertainties. In the case considered, estimates are coherent for 1000-, 300-, and 100-yearly maxima. As expected, the confidence intervals (computed by the delta method, see section 2b) expand as shorter sequences of maxima are used. This also holds for bootstrap and profile likelihood. However, for 50 maxima the profile likelihood confidence intervals become very skewed as opposed to bootstrap or delta method. This clearly indicates poor approximation of normality for the GEV estimators (Coles 2001), underlining the unreliability of the estimates.

Quantile–quantile and return level plots for the above inferences are reported in Fig. 12. These confirm excellent quality for 1000, 300, and 100 maxima, whereas they reveal that something must be wrong for 50. In the quantile–quantile plots (top row of Fig. 12), from left to right increasing departures from the diagonal are apparent especially in the upper tail, whereas the central part of the distribution does not suffer from sample reduction (except in the case of 50 maxima). Analogous effects occur in the upper tail of the return level plots. The main point is that the most delicate part of an extreme value inference is the behavior of the tails. Usually, this is also the aspect one is most interested in. Notice how the black line in the middle of the return level plot for 50 maxima erroneously suggests unboundedness of the return levels (which is only possible for $\xi \geq 0$, see section 2b). Therefore, extrapolations to high levels should be avoided in this case.

We emphasize that the value of T_E, which has just been examined, corresponds to a particularly *bad* inference for 50 yr. An overview, throughout the considered range of T_E, of GEV inference sensitivity to length reduction is summarized in Fig. 13, where the cases of 300-, 100-, and 50-yearly maxima are plotted against 1000. The quality of the fits, of course, generally decreases when using shorter series of maxima. Inference of ξ is particularly sensitive to the length of the series of maxima: the maximal value of the ratios between uncertainty in ξ and value of the corresponding maximum likelihood estimate of ξ is 600%, 1387%, 45%, and

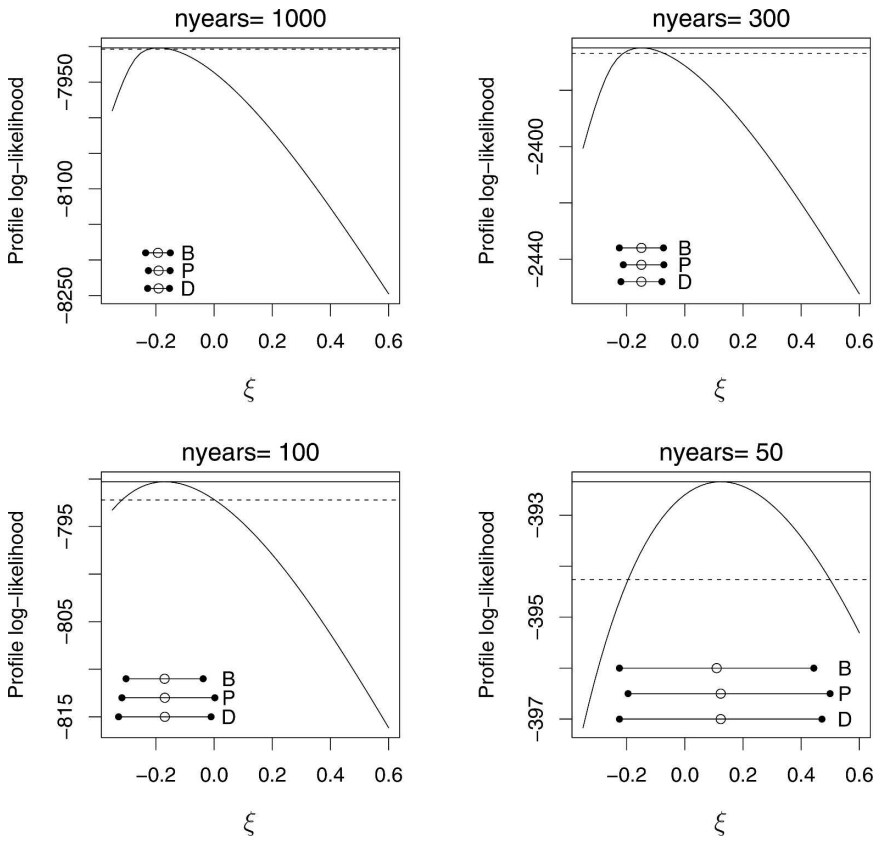

FIG. 10. Profile likelihood plots of ξ for $T_E = 32$, where the number of used yearly maxima is (top left) 1000, (top right) 300, (bottom left) 100, and (bottom right) 50. Confidence intervals are computed by bootstrap, profile likelihood, and observed information matrix (the three stacked segments at the bottom part of the plots labeled by B, P, and D, respectively). The solid and dashed horizontal lines are at the values $\hat{\xi}$ and $\hat{\xi} - 0.5q_{0.95}$, respectively, where $q_{0.95}$ is the 95% quantile of the χ^2 distribution with 1 degree of freedom [cf. Eq. (9)]. Notice the increasing width of the confidence intervals (quite large already for length 100) and the agreement between confidence intervals computed by the three methods, also for the *wrong* (positive) estimate obtained with 50 maxima.

21%, for 50, 100, 300, and 1000 maxima, respectively. The median of those ratios is 48%, 30%, 19%, and 10%, respectively. Taking only 50 maxima yields two positive estimates of ξ (for $T_E = 32$ and 50), which is an unphysical result and overall very large uncertainties: for many values of T_E, confidence bands for ξ include part of the positive axis. The bias in the estimates of ξ induces a significant alteration in those of σ. However, the inferred values of μ display a remarkable insensitivity and the same holds for the 100-yr return levels (Fig. 14), although large uncertainties are obtained for the two cases $T_E = 32$ and 50 that correspond to positive estimates of ξ.

b. *Sensitivity with respect to the extreme events selection procedure: Soft extremes*

We now turn to the second type of inference sensitivity mentioned above, obtained by using so-called soft extremes (Klein Tank and Können 2003) instead of *genuine* extremes. In the present setting, we simulate the usage of soft extremes by considering sequences of maxima over data blocks that correspond to time spans, which are shorter than 1 yr, in particular 0.6, 1.2, and 3 months. In the first two cases, and especially in the first, we are not even sure that the considered maxima are actually uncorrelated, which is the typical situation in real systems. In each case, the number of considered extremes is kept fixed to 1000, so that the difference is only determined by block length.

The net result of using shorter time spans is the introduction of a progressively larger bias in the GEV inferences. The location and shape parameters are systematically underestimated. For the location parameter μ (leftmost column in Fig. 15) the underestimation increases when taking maxima over shorter time spans, but it also increases with T_E. Notice that this is quite

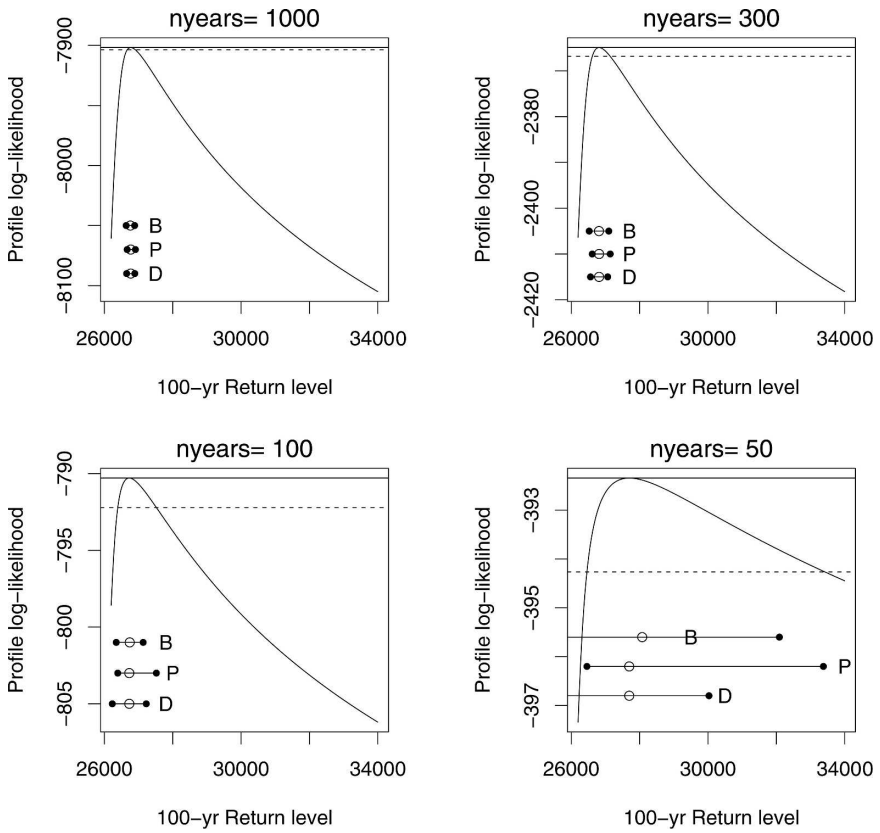

FIG. 11. Profile likelihood plots of the 100-yr return level for $T_E = 32$ for different lengths of the sample of yearly maxima: (top left) 1000-, (top right) 300-, (bottom left) 100-, and (bottom right) 50-yearly maxima. Confidence intervals are computed by bootstrap, profile likelihood, and observed information matrix (the three stacked segments at the bottom part of the plots labeled by B, P, and D, respectively). Notice the increasing width of the confidence intervals and increasing skewness of those obtained by profile likelihood: for 50 maxima the left ends of the B and D confidence intervals lie out of the figure.

different from the effect obtained by reducing the number of maxima (cf. Fig. 13, leftmost column). The sample medians of the relative differences between the estimates of μ for 12 months and those for 3, 1.2, and 0.6 months (where the sample is indexed by the values of T_E for which the estimates are computed) are 3.2%, 5.7%, and 7.5% for 3, 1.2, and 0.6 months, respectively. The underestimation of the 100-yr return levels (see Fig. 16) is a consequence of the underestimation of μ [see the definition in Eq. (10)]. Also notice that the variations in the return levels connected to an increase in T_E are much larger than those induced by usage of either soft extremes or shorter datasets (cf. with Fig. 13, rightmost column). Conversely, the scale parameter σ (second column from the left in Fig. 15) is *largely* overestimated: the sample medians of the relative differences between the estimates of σ are 31%, 59%, and 82% for 3, 1.2, and 0.6 months, respectively. So in our case, taking soft extremes mistakenly suggests an enhanced variability in the extreme values.

Qualitatively, the response of the GEV estimates to the usage of soft extremes is explained by the introduction of much more data in the central part and in the lower tail of the distribution of the selected extreme values. From this fact, the underestimation of μ follows directly. Moreover, since the range of the extreme events distribution gets wider, a larger variability is *artificially* introduced and this is indicated by an overestimated scale parameter σ. Last, the upper tail of the obtained distribution of extremes looks more *squeezed*, given the wider extension at lower values. This corresponds to a more negative value of ξ (cf. the third column from the left in Fig. 15).

c. Sensitivity with respect to the model resolution

In this section we analyze the response of the GEV inferences to variations in the model. In fact, this re-

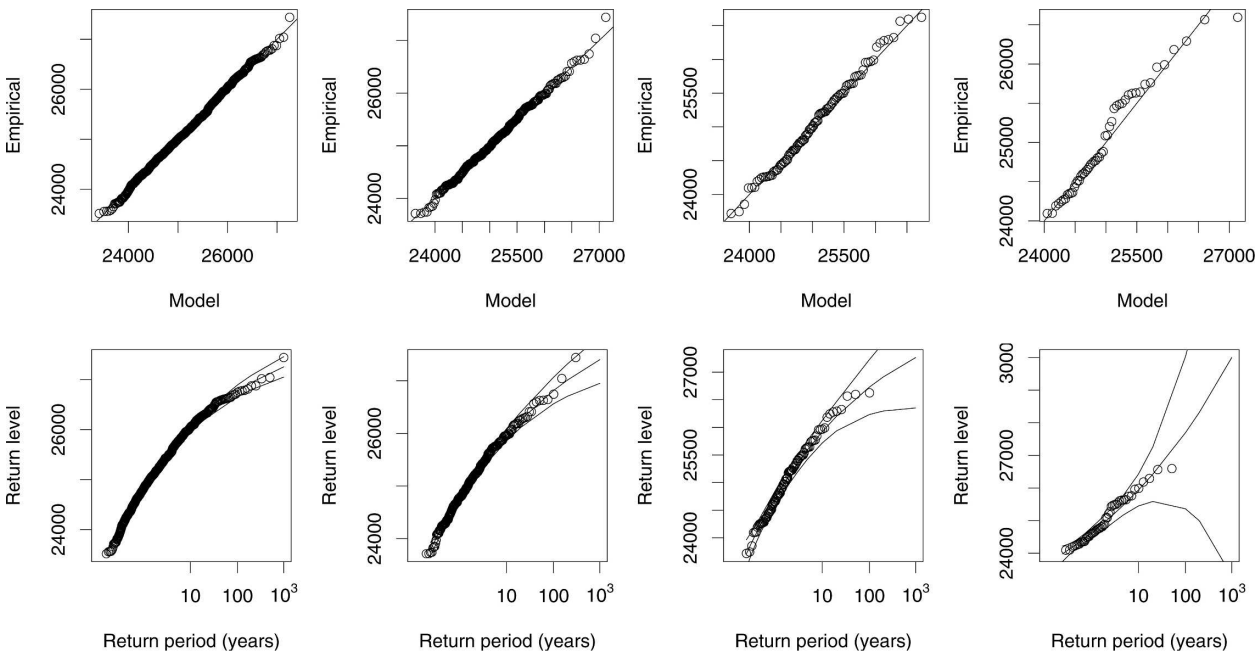

FIG. 12. Diagnostic plots of the GEV inferences for $T_E = 32$. (top row) Quantile–quantile and (bottom row) return level plots, respectively (see section 2b for definitions). Left to right columns are sequences of yearly maxima of the total energy used, having lengths 1000, 300, 100, and 50, respectively. Notice the different scale of the vertical axis in the rightmost return level plot.

sponse is a further aspect of the issues of smoothness and robustness discussed in section 3a, which is of great practical importance: our estimates should not drastically change if the model is slightly altered. Different choices are possible, such as introducing an orography in the bottom layer or changing the lateral boundary conditions. In the present setting, however, we confine ourselves to compare simulations of the baroclinic model performed with a few values of the spectral discretization order J_T [see Eqs. (A11)–(A14) in the appendix].

Time series of the total energy, of length 1000 yr, are computed with the baroclinic model using four different resolutions: $J_T = 8$, 16, 32, and 64 (resolution $J_T = 32$ is used throughout the rest of this paper). In each case the GEV parameters are estimated from sequences of 1000-yearly maxima. The results are compared with each other in Fig. 17. The relative differences of the estimated values of μ between the case $J_T = 64$ and each of the other three cases (top left panel) remain rather small: they are less than 1% and 2% for $J_T = 32$ and 16, respectively, and increase to about 5% for $J_T = 8$. In a similar manner, the 100-yr return level E_{100} of the total energy (top right panel) is not very sensitive to changes in spectral order: for the relative difference of the cases $J_T = 64$ and $J_T = 32$ it is less than 2%. Also the estimates of ξ (bottom right panel) generally agree quite well with all the resolutions considered. More pronounced differences appear in the inferred values of the scale parameter σ (bottom left panel): for $T_E \geq 26$, the estimates obtained with resolutions $J_T = 8$ and 64 are larger than those for $J_T = 16$ and 32.

The estimates for μ and for the 100-yr return level E_{100} closely reflect the behavior of the time-averaged total energy (computed on the same time series from which the yearly maxima are extracted). Considering, to focus ideas, the range $T_E \in [26, 36]$, for each fixed T_E both the inferred values of μ and the time-averaged total energy (not shown) *decrease* as J_T increases. Conversely, there is no simple relation between the sample standard deviation σ_E of the total energy time series and the GEV scale parameter a: for the mentioned values of T_E, the sample standard deviation σ_E decreases for larger J_T (not shown), whereas this is not so for the scale parameter, see above.

Power-law fits of μ and σ as functions of T_E are performed for 1000-yearly maxima of the total energy, where the baroclinic model is run with four different resolutions: $J_T = 8$, 16, 32, and 64. As in section 3a, the range of T_E is divided into two intervals for the fits of μ and into three for σ (in the latter case, no power law is found in the leftmost interval). Remarkable accuracy and coherence of the laws for μ is observed (Table 3).

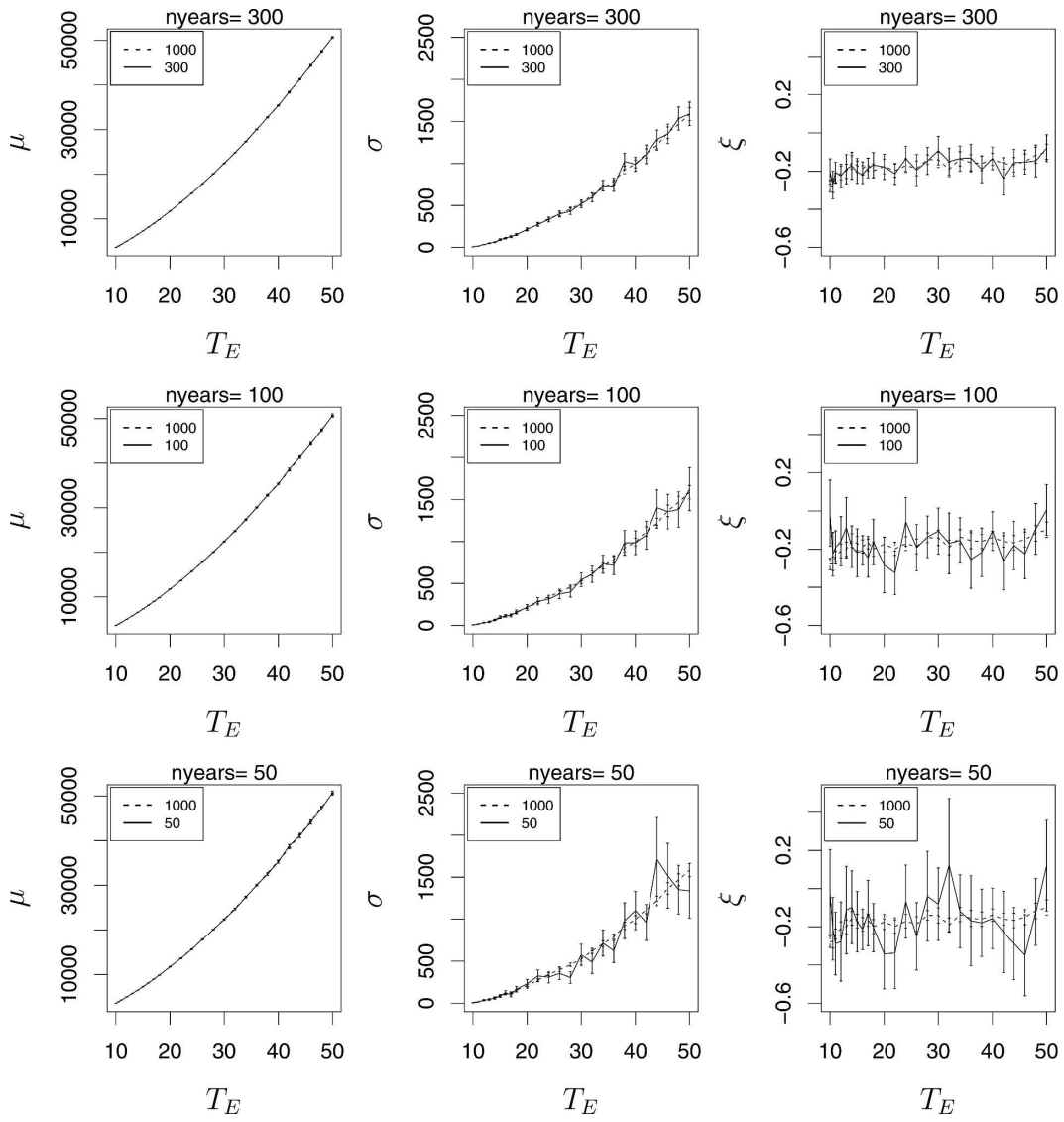

FIG. 13. (top) Maximum likelihood estimates of (left) μ, (middle) σ, and (right) ξ for 1000- and 300-yearly maxima (dashed and solid lines, respectively), with confidence intervals computed by the observed information matrix Eq. (8). (center), (bottom) Same as top but for 100- and 50-yearly maxima, respectively, instead of 300. In the case of 50 maxima, for $T_E = 32$ and 50 the inferred values of ξ are positive (thus completely wrong according to the theoretical expectation, see text) and the uncertainties are very large for σ and ξ.

There is more variability in the power laws for σ (Table 4), although, again, a striking coherence is observed for large T_E.

To summarize, we have observed moderate model sensitivity for the GEV estimates. This inspires confidence in the validity of our results for the class of intermediate complexity models considered in this paper. However, it is to be emphasized that a particularly *stable* observable has been examined here (the total energy) and that only one type of model alteration has been considered, namely a change in the spectral order. More complex models (e.g., general circulation models) might exhibit sensitive behavior with respect to resolution, particularly for phenomena, such as the precipitation, for which the involved spatial scales are small.

5. Summary and conclusions

In this paper we have performed statistical inference of extreme values on time series obtained by a minimal two-level quasigeostrophic model of the atmosphere at midlatitudes. The physical observable used to generate the time series is the total energy of the system and the

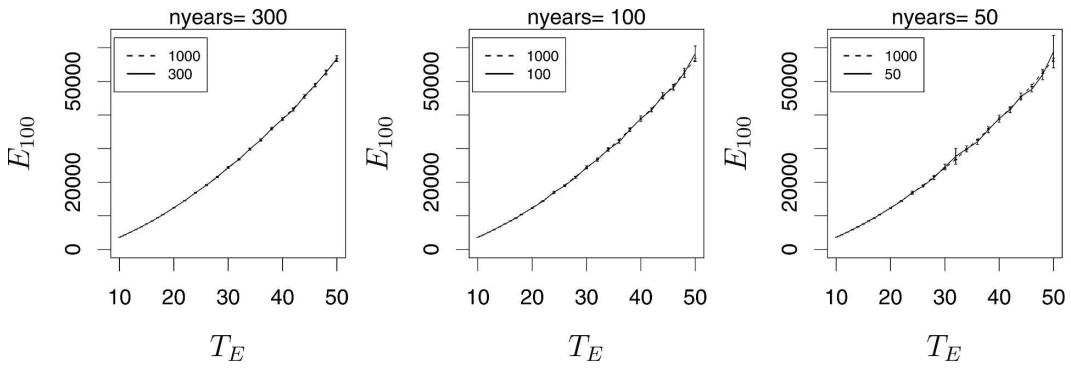

FIG. 14. (left) Maximum likelihood estimates of the 100-yr return level of the total energy as a function of T_E obtained with a sequence of 1000- and 300-yearly maxima (dashed and solid lines, respectively). (middle), (right) Same as left but for where the estimates plotted with solid lines have been computed with 100- and 50-yearly maxima, respectively.

statistical model for the extremes is the generalized extreme value distribution (GEV). Several physically realistic values of the parameter T_E, descriptive of the forced equator-to-pole temperature gradient and responsible for setting the average baroclinicity in the atmospheric model, are examined. In the standard setting, the maxima of the total energy are computed over data blocks with a length of 1 yr and 1000 maxima used as basis for the inference.

A result of the present investigation, which has potential relevance in the field of geophysical fluid dynamics, is the detection of a piecewise smooth dependence of the location and scale GEV parameters (μ, σ) on the macroscopic forcing parameter T_E controlling average baroclinicity. Two distinct power laws, holding in different intervals of T_E, are obtained for both μ and σ as functions of T_E, where the fit for μ is quite accurate. Interestingly, the increase in variability is larger than the increase in average intensity of the extremes. The shape parameter ξ also increases with T_E and it is always negative, as a priori required by the boundedness of the total energy of the system. We conjecture that the dependence of ξ on T_E becomes smooth when much longer time series are considered. The observed smoothness is considered in relation to the results in Lucarini et al. (2005), where analogous scaling laws are found for other dynamical properties, such as Lyapunov exponents and dimension, and physical observables, such as the time–space average of total energy and zonal wind. All of these problems will be explored further in connected work (see Felici et al. 2007).

After the assessment of the goodness-of-fit by means of standard statistical diagnostics, such as return levels and quantile–quantile plots, and computation of confidence intervals by different procedures, we have consistently verified that

- the adopted block length of 1 yr guarantees that the extremes are uncorrelated and genuinely extreme; assessing this property may result more problematic when dealing with real observations because of seasonal modulations, etc;
- the considered length of the series of maxima (1000 data) yields reliable parameter estimates; and
- the GEV inferences are not dramatically affected by structural changes in the atmospheric model adopted in the present work.

The sensitivity of the statistical inference process is first studied with respect to the selection procedure of the maxima: we analyze the effects of reducing either the number of maxima or the length of data blocks over which the maxima are computed.

The first point is checked by repeating the GEV inferences using only 300-, 100-, and 50-yearly maxima. The estimates are coherent for 1000-, 300-, and 100-yearly maxima, but the confidence intervals of the best estimates, not surprisingly, expand as shorter sequences of maxima are used. Moreover, markedly unreliable estimates are obtained when only 50-yearly maxima are considered: the estimated long-term return levels are patently wrong, the uncertainty of the inferred shape parameter ξ is very large, and the best estimate of ξ is positive (i.e., unrealistic) for a few values of T_E.

To address the second point, we have taken maxima over data blocks corresponding to shorter time spans to explore the effects of using soft extremes (Klein Tank and Können 2003). Specifically, the sensitivity of the GEV inferences is analyzed with respect to shortening the length of the data blocks to 3, 1.2, and 0.6 months. The obtained statistics *are* contaminated: an unacceptable bias is introduced for the cases of 1.2 and 0.6 months and still significant (at least for the GEV pa-

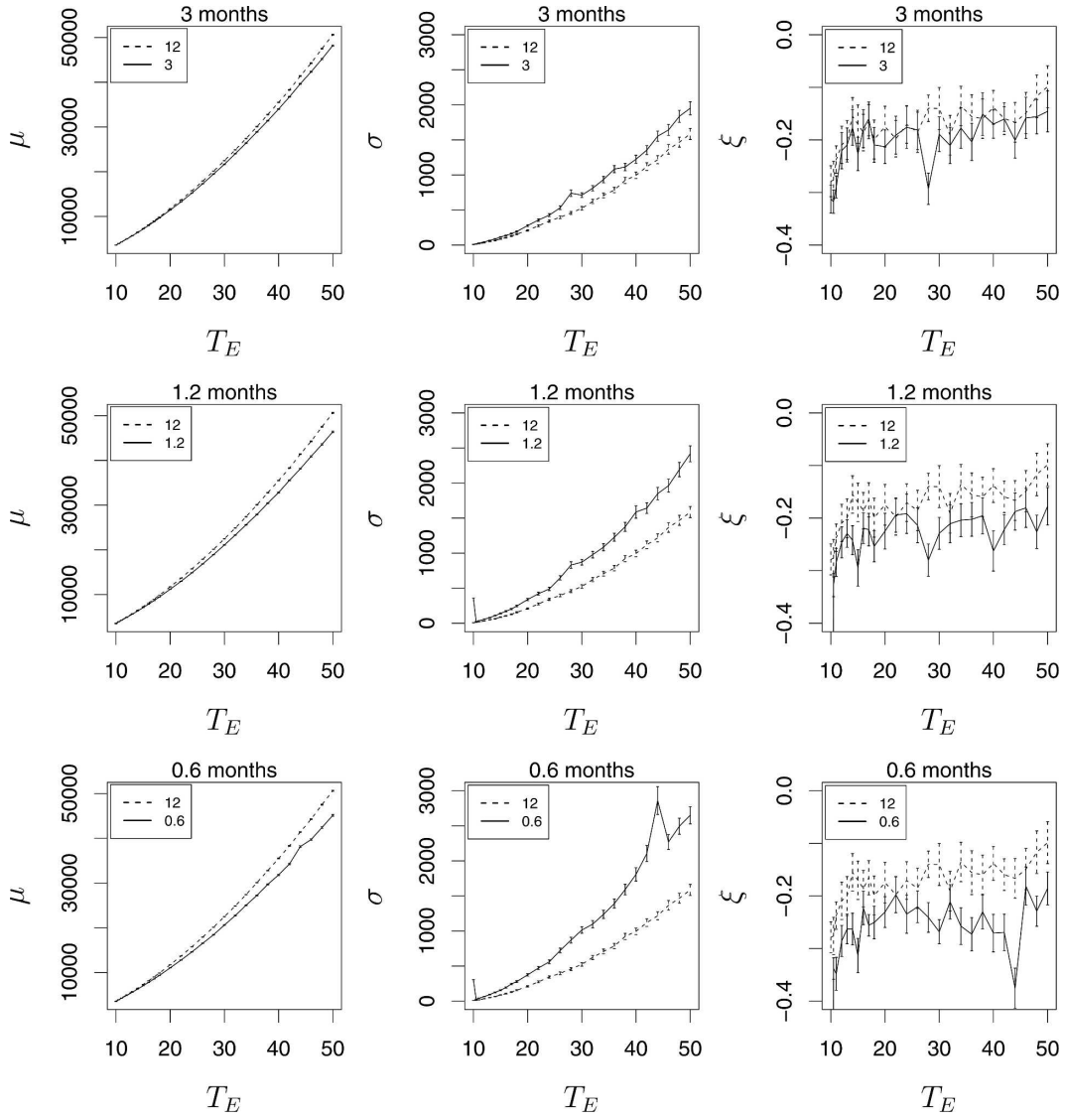

FIG. 15. Inferred values of GEV parameters as a function of T_E in a soft extremes experiment: sequences of 1000 maxima of the total energy time series are used, where the maxima are determined over data blocks corresponding to (top) 3, (center) 1.2, and (bottom) 0.6 months; (left) μ, (middle) σ, and (right) ξ are plotted. Dashed lines are the estimates obtained for the yearly maxima (as in Fig. 5) for reference. Notice how the magnitude of the uncertainties does not significantly depend on the temporal block length.

rameter σ) for 3 months. Moreover, the parameter ξ tends to be underestimated. Taking shorter maxima sequences results in even larger uncertainties, very large for the case of 50-yearly maxima. Physically unrealistic values of ξ may also be obtained.

Finally, issues related to model sensitivity are also explored by varying the (spectral) resolution of the system, and it turns out that the GEV estimates are in general rather robust. Summarizing, to get a good inference *many* maxima are required and they must be genuinely *extreme*, that is, taken over sufficiently large data blocks. Failing to fulfill these requirements may result in affecting the GEV estimates much more seriously than adopting a baroclinic model with lower resolutions. However, it must be emphasized that for more complex models (e.g., general circulation models) and when using spatially localized observables (such as precipitation) high resolution might be essential to get accurate inferences. Moreover, the extremes of localized observables might prove harder to analyze than global observables: a recent study performed with very long integrations of a global atmospheric model shows that when considering a localized observable, such as the air temperature at a given grid point, the convergence to

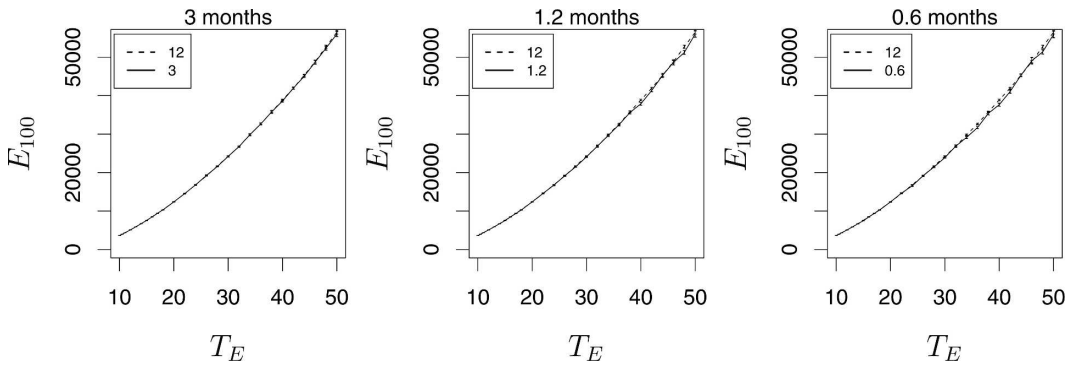

FIG. 16. (left) Maximum likelihood estimates of the 100-yr return level of the total energy as a function of T_E obtained with a sequence of 1000 maxima, where the maxima are computed over data blocks corresponding to 1 yr and 3 months (dashed and solid lines, respectively). (middle), (right) Same as left but for where the estimates plotted with solid lines have been computed using maxima over blocks corresponding to 1.2 and 0.6 months, respectively.

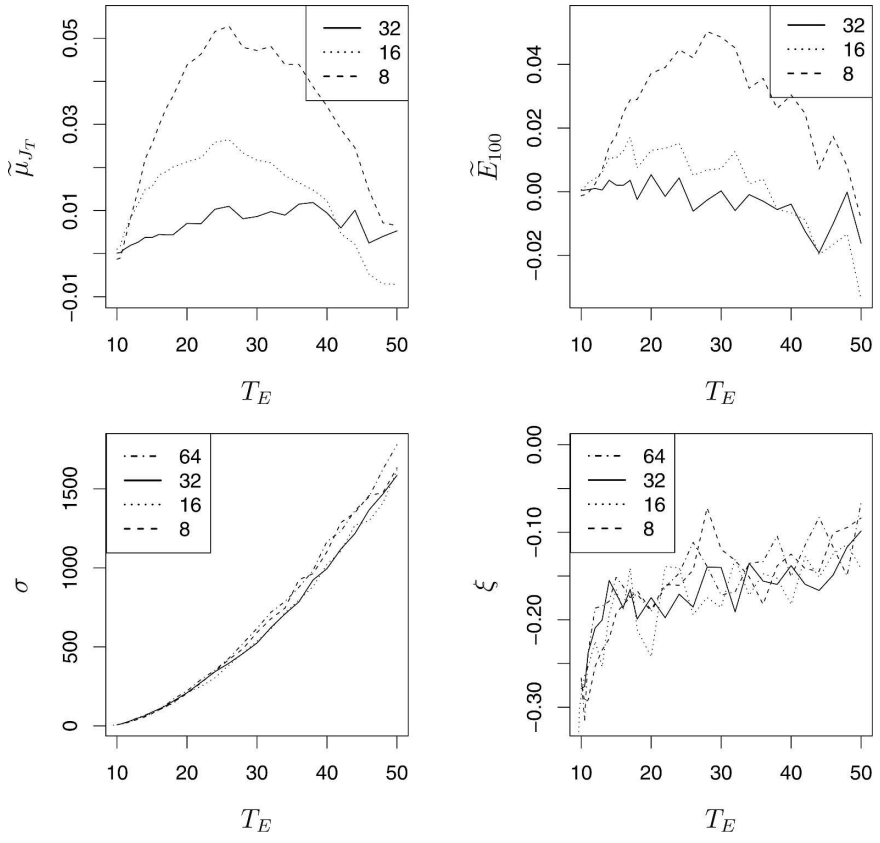

FIG. 17. (top left) Relative difference $\tilde{\mu}_{J_T} = (\mu_{J_T} - \mu_{64})/\mu_{64}$, with respect to the reference case $J_T = 64$, of the maximum likelihood estimates of the GEV parameter μ (vertical axis) for resolutions $J_T = 8$, 16, and 32 (dashed, dotted, and solid lines, respectively) as a function of T_E. (top right) Same as top left but for the relative difference $\tilde{E}_{100,J_T} = (E_{100,J_T} - E_{100,64})/E_{100,64}$ of the 100-yr return level of the energy. (bottom left) Estimates of σ as a function of T_E for the cases $J_T = 8$, 16, 32, and 64 (dashed, dotted, solid, and dashed–dotted lines, respectively); (bottom right) Same as bottom left but for ξ. Sequences of 1000 maxima are used for all estimates.

TABLE 3. Power-law fits of the location parameter μ as a function of T_E of the form $\mu \propto T_E^{\gamma}$. Here J_T indicates the spectral resolution (number of Fourier modes) of the baroclinic model and T_E^b is the value of T_E dividing the two considered intervals (see text for details).

J_t	$\gamma_{\mu,1}$	T_E^b	$\gamma_{\mu,2}$
64	1.7346 ± 0.0008	15	1.6027 ± 0.0005
32	1.7310 ± 0.0007	18	1.6019 ± 0.0005
16	1.7027 ± 0.0007	18	1.5982 ± 0.0007
8	1.6794 ± 0.0006	22	1.5977 ± 0.0011

TABLE 4. Same as in Table 3 but for the scale parameter $\sigma \propto T_E^{\gamma}$. The interval $[T_E^{b1}, T_E^{b2}]$ is the range of validity of the first power law having exponent γ_1. The point dividing the two considered intervals is T_E^{b2}. No power law is detected for $T_E < T_E^{b1}$.

J_T	T_E^{b1}	$\gamma_{\sigma,1}$	T_E^{b2}	$\gamma_{\sigma,2}$
64	18	2.514 ± 0.046	32	2.067 ± 0.055
32	15	3.011 ± 0.076	22	2.140 ± 0.025
16	15	2.821 ± 0.045	26	2.150 ± 0.033
8	17	2.675 ± 0.065	26	2.149 ± 0.033

the GEV family of the empirical distribution of the block maxima is very slow and multiannual maxima have to be considered, maybe because of spatial correlation effects (Vannitsem 2007).

If we consider the additional complications present in observed data or in more realistic representations of the natural processes (e.g., taking into account seasonal modulations), it is apparent that a reliable estimation of the uncertainties in the extremes of a given meteoclimatic variable is crucial. In the present case, the most robust statistical properties turn out to be the return levels, which are also the most relevant for applications. We maintain that a rigorous and well-defined framework for the statistical analysis of extremes of observed data, such as that provided by the GEV theory, is necessary to study the past and present climate and to characterize its variations.

We conclude by highlighting that the parameterization of physical observables with respect to an external forcing is indeed a rather general and difficult problem in the dynamical analysis of a physical system. Existence of a unique Sinai–Ruelle–Bowen (SRB) measure is *required* in order to rigorously associate a stationary stochastic process to the dynamical evolution law. However, even if an SRB measure exists and is unique, there is typically no explicit expression in terms of the system's equations and parameters (Eckmann and Ruelle 1985). Both our conjecture about the existence of a unique SRB measure for the analyzed system and the observed smooth dependence on T_E of all the considered statistical and dynamical indicators are coherent with the theory proposed by Gallavotti and Cohen (Gallavotti and Cohen 1995a,b; Gallavotti 1996; Cohen and Gallavotti 1999). Indeed, the chaotic hypothesis implies that, for the purpose of computing macroscopic quantities, the attractor of the system behaves as though it were structurally stable with respect to changes in the external parameters. In this respect, the simplicity and the universality of the GEV model can be exploited to characterize chaotic systems by focusing on extreme values of suitable time series, rather than examining the distribution of all states visited by the system in phase space (i.e., the SRB measure). Different model variants (both in boundary conditions and in model structure) and other observables will be considered in future research.

Acknowledgments. The authors wish to thank Stefano Pittalis and Nazario Tartaglione for useful conversations and two anonymous referees for valuable suggestions. This work has been supported by MIUR PRIN Grant "Gli estremi meteo-climatici nell'area mediterranea: Proprietà statistiche e dinamiche," Italy, 2003.

APPENDIX

A Model for the Midlatitude Atmospheric Circulation

As mentioned in the introduction, the stochastic generator of the energy time series used in this paper is a model for the baroclinic jet at midlatitudes. The system is relaxed toward a prescribed north–south temperature profile, where the gradient is controlled by the parameter T_E. In fact, the parameter T_E controls average baroclinicity of the system and is used to study the relation with extreme values of the energy time series. Many dynamical properties of the model depending on T_E have been analyzed before the analysis of extreme values presented in this paper, providing a sort of road map. See Speranza and Malguzzi (1988); Malguzzi et al. (1990); and Lucarini et al. (2005) for a detailed derivation of the model and for discussion on the physics involved. In this section, we confine ourselves to a brief sketch.

Starting point for the construction of the model is the two-level quasigeostrophic equation:

$$\frac{\partial}{\partial t}\Delta_H\tau - \frac{2}{H_2^2}\frac{\partial}{\partial t}\tau + J\left(\tau, \Delta_H\phi + \beta y + \frac{2}{H_2^2}\phi\right) + J(\phi, \Delta_H\tau)$$
$$= \frac{2\upsilon_E}{H_2^2}\Delta_H(\phi - \tau) - \frac{2\kappa}{H_2^2}\Delta_H\tau + \frac{2\upsilon_N}{H_2^2}(\tau - \tau^*), \quad \text{and} \tag{A1}$$

$$\frac{\partial}{\partial t}\Delta_H\phi + J(\phi, \Delta_H\phi + \beta y) + J(\tau, \Delta_H\tau) = -\frac{2\upsilon_E}{H_2^2}\Delta_H(\phi - \tau). \quad \text{(A2)}$$

Here τ and ϕ are the baroclinic and barotropic components, respectively, of the streamfunction ψ_1 and ψ_3 at the two levels

$$\tau = \frac{1}{2}(\psi_1 - \psi_3) \quad \text{and} \quad \phi = \frac{1}{2}(\psi_1 + \psi_3), \quad \text{(A3)}$$

where Δ_H is the horizontal Laplacian, $1/H_2^2$ is the Froude number, β is the gradient of the Coriolis parameter, and υ_E, κ, and υ_N parameterize the Ekman pumping at the lower surface, the heat diffusion, and the Newtonian cooling, respectively.

The system is driven for the baroclinic component by the term in $(\tau - \tau^*)$ in Eq. (A1), which forces a relaxation to the radiative equilibrium τ^* with a characteristic time scale of $1/\upsilon_N$. We take

$$\tau^* = \frac{R}{f_0}\frac{T_E}{4}\cos\left(\frac{\pi y}{L_y}\right), \quad \text{(A4)}$$

so that T_E is the forced temperature difference between the low- and high-latitude border of the domain. In this sense, the parameter T_E is responsible for average baroclinicity of the system and is the control parameter we vary to test changes in the extreme value statistics.

The fields ϕ and τ are expanded in Fourier series in the longitudinal direction x. Moreover, in order to avoid wave–wave nonlinear interactions, only the terms of order $n = 1$ and $n = 6$ are retained (see Lucarini et al. 2005 for details). These yield

$$\phi(x, y, t) = -\int_0^y U(z, t)\, dz + A\exp(i\chi x) + \text{c.c.} \quad \text{and} \quad \text{(A5)}$$

$$\tau(x, y, t) = -\int_0^y m(z, t)\, dz + B\exp(i\chi x) + \text{c.c.} \quad \text{(A6)}$$

By substitution into Eqs. (A1)–(A2), one obtains

$$\dot{A}_{yy} - \chi^2\dot{A} + \left(i\chi U + \frac{2\upsilon_E}{H_2^2}\right)A_{yy} - \left(i\chi^3 U + i\chi U_{yy} + \frac{2\upsilon_E}{H_2^2}\chi^2 - i\chi\beta\right)A + \left(i\chi m - \frac{2\upsilon_E}{H_2^2}\right)B_{yy} - \left(i\chi^3 m + i\chi m_{yy} - \frac{2\upsilon_E}{H_2^2}\chi^2\right)B = 0, \quad \text{(A7)}$$

$$\dot{B}_{yy} - \chi^2\dot{B} - \frac{2}{H_2^2}\dot{B} + \left(i\chi U + \frac{2\upsilon_E}{H_2^2} + \frac{2\kappa}{H_2^2}\right)B_{yy} - \left(i\chi^3 U + i\chi U_{yy} + \frac{2\upsilon_E}{H_2^2}\chi^2 - i\chi\beta + \frac{2\kappa}{H_2^2}\chi^2 + \frac{2\upsilon_N}{H_2^2} + \frac{2}{H_2^2}i\chi U\right)B + \left(i\chi m - \frac{2\upsilon_E}{H_2^2}\right)A_{yy} - \left(i\chi^3 m + i\chi m_{yy} - \frac{2\upsilon_E}{H_2^2}\chi^2 - \frac{2}{H_2^2}i\chi m\right)A = 0, \quad \text{(A8)}$$

$$\dot{U} + \frac{2\upsilon_E}{H_2^2}(U - m) + 2\chi\,\text{Im}(AA_{yy}^* + BB_{yy}^*) = 0, \quad \text{and} \quad \text{(A9)}$$

$$\dot{m}_{yy} - \frac{2}{H_2^2}\dot{m} + \frac{2\kappa}{H_2^2}m_{yy} - \frac{2\upsilon_E}{H_2^2}(U - m)_{yy} - \frac{2\upsilon_N}{H_2^2}(m - m^*) + \frac{4}{H_2^2}\chi\,\text{Im}(A^*B)_{yy} + 2\chi\,\text{Im}(AB_y^* + BA_y^*)_{yyy} = 0, \quad \text{(A10)}$$

where the dot indicates time differentiation and A^* denotes the complex conjugate of A. This is a set of six equations for the real fields A^1, A^2, B^1, B^2, U, and m, where A^1 and A^2 are the real and imaginary parts of A and similarly for B. Rigid walls are taken as boundaries at $y = 0$, L_y, so that all fields have vanishing boundary conditions.

A system of ordinary differential equations is obtained from Eqs. (A7) to (A10) by means of a pseudospectral (collocation) projection involving a Fourier half-sine expansion of the fields of the form

$$A^i = \sum_{j=1}^{J_T} A_j^i \sin\left(\frac{\pi j y}{L_y}\right), \quad i = 1, 2, \quad \text{(A11)}$$

$$B^i = \sum_{j=1}^{J_T} B_j^i \sin\left(\frac{\pi j y}{L_y}\right), \quad i = 1, 2, \quad \text{(A12)}$$

$$U = \sum_{j=1}^{J_T} U_j \sin\left(\frac{\pi j y}{L_y}\right), \quad \text{and} \quad \text{(A13)}$$

$$m = \sum_{j=1}^{J_T} m_j \sin\left(\frac{\pi j y}{L_y}\right). \quad \text{(A14)}$$

The resulting system is the generator of the time series used in this paper for extreme value analysis. In particular, as *observable* (i.e., as function of the state space yielding the time series) we choose the total energy $E(t)$ of the system, obtained by integration in the (x, y) domain of the energy density:

$$e(x, y, t) = \frac{\delta p}{g}\left[\frac{1}{2}(\nabla\psi_1)^2 + \frac{1}{2}(\nabla\psi_3)^2 + \frac{1}{2H_2^2}(\psi_1 - \psi_3)^2\right]. \tag{A15}$$

Here the factor $\delta p/g$ is the mass per unit surface in each level; the first two terms inside the brackets describe the kinetic energy and the last term describes the potential energy. We emphasize that in Eq. (A15) the potential energy term is half of what is reported in Pedlosky (1987), which contains a trivial algebraic mistake.

It turns out that the order $J_T = 32$ in Eqs. (A11)–(A14) is sufficiently high to have an earthlike chaotic regime characterized by intermediate dimensionality in suitable ranges of the parameter T_E. By chaotic, we mean that the dynamics take place on a strange attractor with internally generated noise. By earthlike we mean that the time-dependent Fourier coefficients in Eqs. (A11)–(A14), as well as the total energy and mean zonal wind, have unimodal probability densities. The mentioned chaotic range is $T_E > T_E^{\text{crit}}$, where $T_E^{\text{crit}} = 8.75$ approximately. For lower values of T_E, the Hadley equilibrium (stationary solution) is stable and is therefore the unique attractor. See Speranza and Malguzzi (1988); Malguzzi et al. (1990); and Lucarini et al. (2005) for a complete discussion. Throughout this work, we consider $J_t = 8$, 16, 32, and 64 and the considered parameter range is $10 \leq T_E \leq 50$ with integer steps of 2.

REFERENCES

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature,* **419,** 224–232.

Castillo, E., 1988: *Extreme Value Theory in Engineering.* Academic Press, 389 pp.

Cohen, E. G. D., and G. Gallavotti, 1999: Note on two theorems in nonequilibrium statistical mechanics. *J. Stat. Phys.,* **96,** 1343–1349.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, 208 pp.

Eckmann, J.-P., and D. Ruelle, 1985: Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.,* **57,** 617–655.

Embrechts, P., C. Kluppelberg, and T. Mikosch, 1997: *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag, 645 pp.

Felici, M., V. Lucarini, A. Speranza, and R. Vitolo, 2007: Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part II: Trend detection and assessment. *J. Atmos. Sci.,* **64,** 2159–2175.

Fisher, R. A., and L. H. C. Tippett, 1928: Limiting forms of the frequency distribution of the largest or smallest number of a sample. *Proc. Cambridge Philos. Soc.,* **24,** 180–190.

Galambos, J., 1978: *The Asymptotic Theory of Extreme Order Statistics.* Wiley, 352 pp.

Gallavotti, G., 1996: Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem. *J. Stat. Phys.,* **84,** 899–926.

——, and E. G. D. Cohen, 1995a: Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.,* **74,** 2694–2697.

——, and ——, 1995b: Dynamical ensembles in stationary states. *J. Stat. Phys.,* **80,** 931–970.

Gnedenko, B. V., 1943: Sur la distribution limite du terme maximum d'une série aléatorie. *Ann. Math.,* **44,** 423–453.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis.* Cambridge University Press, 881 pp.

Ihaka, R., and R. Gentleman, 1996: R: A language for data analysis and graphics. *J. Comput. Graph. Stat.,* **5,** 299–314.

Jenkinson, A. F., 1955: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. Roy. Meteor. Soc.,* **81,** 158–171.

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Amer. Meteor. Soc.,* **79,** 231–241.

——, ——, D. R. Easterling, and R. G. Quayle, 1996: Indices of climate change for the United States. *Bull. Amer. Meteor. Soc.,* **77,** 279–292.

Katz, R. W., and B. G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change,* **21,** 289–302.

——, M. B. Parlange, and P. Naveau, 2002: Statistics of extremes in hydrology. *Adv. Water Res.,* **25,** 1287–1304.

Kharin, V. V., and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J. Climate,* **13,** 3760–3788.

Klein Tank, A. M. G., and G. P. Können, 2003: Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J. Climate,* **16,** 3665–3680.

Kunkel, K. E., R. A. Pielke Jr., and S. A. Changnon, 1999: Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: A review. *Bull. Amer. Meteor. Soc.,* **80,** 1077–1098.

Lavagnini, A., A. M. Sempreviva, C. Transerici, C. Accadia, M. Casaioli, S. Mariani, and A. Speranza, 2006: Offshore wind climatology over the Mediterranean basin. *Wind Energy,* **9,** 251–266.

Leadbetter, M. R., G. Lindgren, and H. Rootzén, 1983: *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, 336 pp.

Lionello, P., F. Dalan, and E. Elvini, 2002: Cyclones in the Mediterranean Region: The present and the doubled CO_2 climate scenarios. *Climate Res.,* **22,** 147–159.

Lorenz, E. N., 1967: *The Nature and Theory of the General Circulation of the Atmosphere.* World Meteorological Organization, 161 pp.

Lucarini, V., 2002: Towards a definition of climate science. *Int. J. Environ. Pollut.,* **18,** 409–414.

——, A. Speranza, and R. Vitolo, cited 2005: Physical and mathematical properties of a quasi-geostrophic model of intermediate complexity of the mid-latitudes atmospheric circulation. [Available online at http://arxiv.org/abs/physics/0511208.]

Malguzzi, P., A. Trevisan, and A. Speranza, 1990: Statistic and predictability for an intermediate dimensionality model of the baroclinc jet. *Ann. Geophys.,* **8,** 29–36.

Morrison, J. E., and J. A. Smith, 2002: Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resour. Res.,* **38,** 1305, doi:10.1029/2001WR000502.

Mortensen, N. G., L. Landberg, I. Troen, E. L. Petersen, 1993: Wind Atlas Analysis and Application Program (WAsP): User's guide. Risø-I-666 (EN), Risø National Laboratory, Roskilde, Denmark.

Nordhaus, W. D., 1994: *Managing the Global Commons: The Economics of Climate Change.* MIT Press, 213 pp.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* 2d ed. Springer-Verlag, 710 pp.

Perrin, O., H. Rootzén, and R. Taesler, 2006: A discussion of statistical methods used to estimate extreme wind speeds. *Theor. Appl. Climatol.,* **85,** 203–215.

Phillips, N. A., 1954: Energy transfonnations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus,* **6,** 273–286.

Rootzén, H., and N. Tajvidi, 2001: Can losses caused by wind storms be predicted from meteorological observations? *Scand. Actuarial J.,* **2001,** 162–175.

Smith, R. L., cited 2006: Trends in rainfall extremes. [Available online at http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html.]

Speranza, A., and P. Malguzzi, 1988: The statistical properties of a zonal jet in a baroclinic atmosphere: A semilinear approach. Part I: Quasi-geostrophic, two-layer model atmosphere. *J. Atmos. Sci.,* **45,** 3046–3062.

Vannitsem, S., 2007: Statistical properties of the temperature maxima in an intermediate order Quasi-Geostrophic model. *Tellus,* **59A,** 80–95.

Watson, R. T., and Coauthors, Eds., 2001: *Climate Change 2001: Synthesis Report.* Cambridge University Press, 398 pp.

Zhang, X., F. W. Zwiers, and G. Li, 2004: Monte Carlo experiments on the detection of trends in extreme values. *J. Climate,* **17,** 1945–1952.

Zwiers, F. W., and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling. *J. Climate,* **11,** 2200–2222.