



**HAL**  
open science

## Modeling Uncertainties in Lichenometry Studies

Philippe Naveau, Vincent Jomelli, Daniel Cooley, Grancher Delphine, Antoine Rabatel

► **To cite this version:**

Philippe Naveau, Vincent Jomelli, Daniel Cooley, Grancher Delphine, Antoine Rabatel. Modeling Uncertainties in Lichenometry Studies. *Arctic, Antarctic, and Alpine Research*, 2007, 39 (2), pp.277-285. 10.1657/1523-0430(2007)39[277:MUJLS]2.0.CO;2 . hal-03191580

**HAL Id: hal-03191580**

**<https://hal.science/hal-03191580v1>**

Submitted on 7 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Modeling Uncertainties in Lichenometry Studies**

Authors: Naveau, Philippe, Jomelli, Vincent, Cooley, Daniel, Delphine, Grancher, and Rabatel, Antoine

Source: Arctic, Antarctic, and Alpine Research, 39(2) : 277-285

Published By: Institute of Arctic and Alpine Research (INSTAAR),  
University of Colorado

URL: [https://doi.org/10.1657/1523-0430\(2007\)39\[277:MUJLS\]2.0.CO;2](https://doi.org/10.1657/1523-0430(2007)39[277:MUJLS]2.0.CO;2)

---

BioOne Complete ([complete.BioOne.org](https://complete.BioOne.org)) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at [www.bioone.org/terms-of-use](https://www.bioone.org/terms-of-use).

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

---

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

# Modeling Uncertainties in Lichenometry Studies

Philippe Naveau\*§

Vincent Jomelli†

Daniel Cooley\*

Grancher Delphine† and

Antoine Rabatel‡

\*Department of Applied Mathematics,  
University of Colorado, Boulder,  
Colorado 80309, U.S.A.

†Laboratoire de Géographie Physique,  
CNRS, Meudon, France.

‡Laboratoire de Glaciologie et  
Géophysique de l'Environnement, IRD,  
Grenoble, France.

§Corresponding author. Present  
address: Laboratoire des Sciences du  
Climat et de l'Environnement,  
IPSL-CNRS, Orme des Merisiers, 91191  
Gif-sur-Yvette, France.  
naveau@lsce.cnrs-gif.fr

## Abstract

To date glacial and periglacial landforms, lichenometry is a valuable method but, to improve efficiency, the estimated surface dates derived from traditional methods need to be more accurate. In other words, the statistical uncertainty associated with inferred dates has to be reduced. How to perform such a reduction is the main question that we will address in this paper. An interdisciplinary approach (lichenometry and statistics) allows reduction in the main sources of uncertainty: lichen diameters and their associated ages. Around 2600 lichen measurements collected on moraines from the Charquini glacier in Bolivia (Cordillera Real) are used to illustrate the advantages of our approach over past studies.

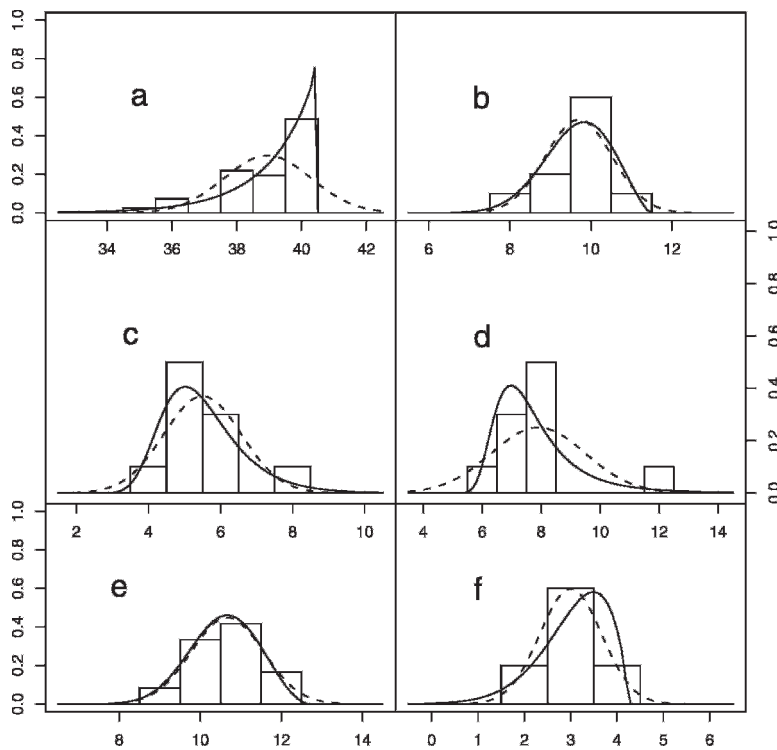
As for any statistical estimation procedure, the error analysis in lichenometry is directly linked to the type of observations and the statistical model used to represent accurately these data. The attribute of lichenometry studies is that the measurements are not averages but maxima; only the largest lichen diameters provide information about the surface ages. To take this characteristic into account, we propose a novel statistical way to model maximum lichen diameters. Our model, based on the extreme value theory, allows us to compute small confidence intervals for the inferred surface ages. In addition, it offers three other advantages: (1) a global statistical model, as all our data (dated surfaces and all lichen maximum diameters) are represented with a unique function; (2) a mathematical framework within which the maximum lichen distribution is derived from a statistical theory; and (3) flexibility, as different types of growing curves can be investigated.

## Introduction

Lichenometry has principally been used to date periglacial and glacial landforms. Since the pioneering work of Beschel (1961), this technique has been applied to a large variety of environments. It is especially well suited for arctic and alpine regions because other dating methods are either difficult to implement or even fail at these high altitudes. For example, the sparsity of vegetation near glaciers makes the use of dendrochronology problematic. Although lichenometry may be used to date old surfaces (e.g. the late Holocene in South America; Rodbell, 1992), it is best suited to analyze recent centuries for which classical  $^{14}\text{C}$  dating techniques are tainted with a low precision.

The basic premise of lichenometry is that the diameter of the largest thalli growing on a surface is proportional to the length of time that the surface has been exposed to colonization and growth in a specific environmental context. Hence, if one can establish a temporal link between lichen sizes and their ages, i.e. construct a growth curve, it is then possible to date moraines. However, this relationship between the lichen size and its age varies with the characteristics of the surroundings. Lichen growth depends on the climate (Benedict, 1967, 1990, 1991), the lithology (Rodbell, 1992), as well as the orientation of the colonized surface (Pentecost, 1979). The most popular approach that involves measurements of the largest lichens is assuming that the largest individual lichens are among the first to colonize a surface. Although there is a common agreement among geomorphologists about the basic concepts of this method, there is still controversy concerning its implementation, e.g. the number of lichens sampled, the design of the experiment, and the choice of a robust statistic.

Recently, McCarroll (1994) proposed a new strategy where only the largest lichen on a block is selected. This procedure is then repeated many times on different blocks. However, McCarroll incorrectly assumed that the largest lichen sizes are near normally distributed (see figures 2 and 7 in McCarroll, 1994). This assumption was only based on a visual inspection of histograms, and neither statistical tests were provided to confirm this choice nor mathematical concepts employed to justify this approach. The latter point is important because probability theory dedicated to extreme values (largest lichen diameters belonging to this category) dictates that the distribution of maxima cannot be normal (Fig. 1) but instead must follow a specific distribution (the Generalized Extreme Value distribution), whenever the sample size is large enough. This mathematical result has been used in many fields since the pioneering work of Gnedenko (1943) but has rarely been applied to lichenometry (e.g. Bull and Brandon, 1998; Karlen and Black, 2002). (A detailed description of this theory, as well as a list of references, will be given in the section *MODELING MAXIMUM DIAMETERS*.) An immediate consequence of this mathematical constraint is that past computations of confidence intervals based on the assumption of normality are at best questionable and at worst flawed. To illustrate the importance of the appropriate choice for the distribution of maxima, we plot in Figure 1 the histogram of the maximum lichen diameters from the Charquini glacier in Bolivia (the same result has been obtained for different moraines and regions). From the histogram shapes, it is very clear that the distributions of maxima are rarely symmetric, and the fit by a Gaussian density distribution (dotted line) is too restrictive to represent accurately maximum lichen sizes (e.g. Fig. 1a). In comparison, the fit by a Generalized Extreme Value



**FIGURE 1.** Distribution of maxima lichen diameters (mm) on six moraines of the Charquini south glacier. The black boxes corresponds to the histogram of the data. The dotted line shows the fit by a Gaussian density distribution. The solid line indicates the fit by a Generalized Extreme Value (GEV) density distribution. The number of lichen per panel is about 20.

(GEV) density distribution (solid line) that is presented in detail in *MODELING MAXIMUM DIAMETERS* provides a much better agreement with the data, because the GEV spans a much wider range of distribution shape and it is specially adapted to deal with maxima. In the past, other authors have noticed this discrepancy between the maximum distribution and a Gaussian fit. Bull and Brandon (1998) and Karlen and Black (2002) made use of the Gumbel distribution to model the maximum diameters density, but these authors did not take full advantage of the extreme value theory. The Gumbel density is only a special case of the GEV density which is a larger class of distribution (see *MODELING MAXIMUM DIAMETERS*). In addition, no statistical model based on the Gumbel distribution was implemented to derive estimation errors and confidence intervals. A few past studies (e.g. O'Neal and Schoenenberger, 2003) attempted to associate confidence intervals and/or p-values with specific growth curves. However, the hypothesis to derive these confidence intervals, when implementing linear or nonlinear regression models, was formulated as: *maximum diameter = l(age) + noise*, where *l* is the growth curve and the noise Gaussian. This last assumption, as seen in Figure 1, is not satisfied.

One of our main goals in this paper is to introduce a sound statistical model based on extreme value theory, which is more appropriate to model the distribution of the raw material of lichenometry: the largest lichen diameters.

This paper is organized as follows. In the next section, a model specifically tailored for analyzing maxima and based on the statistical extreme values theory is presented. A full description of the field site (Cordillera Real, Bolivia) can be found in the section *AREA OF INTEREST*. The observation collection process is detailed in the section *FIELD WORK*. This field study is a part of a larger French program (Great Ice) dedicated to the study of climate and environmental variability in the Andes and during the Little Ice Age (Solomina et al., 2007). To our knowledge, no previous lichenometry studies have been undertaken in Bolivia, and consequently there is no growth curve for this region. The section Conclusions and Future Work summarizes our results.

## A General Statistical Model for Lichen Diameters Analysis

To develop a statistical model for analyzing lichen diameters, we first need to identify the main sources of randomness. In our approach, we focus on two main sources of uncertainty: the error in modeling lichen diameters, and the error of associating a date with a moraine. Consequently, we model a bivariate random vector, say  $(X_i, T_i)$ , the date of the *i*-th moraine  $T_i$  and the lichen diameters associated to this moraine  $X_i$ , where the moraine index *i* varies from 1 to *m*, the total number of moraines. With these notations,  $X_i = \{X_{i,1}, \dots, X_{i,n_i}\}$  is a vector whose length  $n_i$  corresponds to the number of sampled blocks on the *i*-th moraine. Apart from a few cases for which the date of the moraine is known, the variable  $T_i$  is not observed and has to be estimated from the lichen diameters. Deriving this date is one of the most important questions in the analysis of lichen diameters. To solve this problem, we propose a statistical model that has the advantage of being global, i.e. the lichen information (the measured diameters and the moraine dates when available) is pooled into one data set. In past studies, two distinct data sets were considered; lichens with known dates were analyzed separately from lichens with unknown dates (e.g. McCarroll, 1994). Then, these two groups were linked by the growth curve. We argue that all parameters describing the temporal evolution of lichen diameters should be estimated simultaneously. The two main reasons for such a position are the following: (1) Separating the lichens into two groups is statistically arbitrary because it is typically assumed that the distribution of lichen diameters comes from the same family of distribution (otherwise it would not be possible to make inference from dated moraines to undated ones). (2) Past two-step procedures have the disadvantage of propagating the error generated from the first step into the second stage of the estimation of dates. In comparison, a global pooling of all data allows us to maximize a global criteria to estimate the parameters of our statistical model, and consequently it reduces the estimation error.

Another difference with past studies is that the dates  $T_i$  are not necessarily considered as real numbers but are viewed as

random variables, i.e. realizations from a common distribution function. This approach allows us to measure the second source of uncertainty, the error in dating the lichens. For example, carbon dating does not provide a unique date but instead proposes a range of values for the moraine date. To the best of our knowledge, the integration of these two uncertainties into a global framework is novel in most lichenometry studies. Note that capital letters (e.g.,  $X$  and  $T$ ) are used to denote random variables, and lowercase letters (e.g.,  $x$  and  $t$ ) correspond to realizations, i.e. specific outputs for the random variables.

### MODELING MAXIMUM DIAMETERS

Extreme value theory has been applied to a variety of problems in finance (Embrechts et al., 1997), hydrology (Katz et al, 2002), and ecology (Gaines and Denny, 1993). In recent years, the statistical methodology for extremes has also been used for a wide range of problems in climate studies. Extreme climate events due to their large economic and human impacts are becoming more and more studied in geophysics. For example, Kharin and Zwiers (2000) examined changes in climate extremes from coupled atmosphere-ocean GCM outputs, while Naveau and Moncrieff (2003) investigated the role of extremes on upward mass fluxes. To learn more about extreme value theory, the book by Leadbetter et al. (1983) is recommended. For more recent and applied work, we refer to Coles (2001), Smith (1992), Naveau et al. (2005), and Reiss (1997). But what is exactly extreme value theory? Its basic principles come from an asymptotic probabilistic result. The original focus of the theory was to describe the behavior of maxima when the sample size became larger and larger. To better understand this approach, we will first recall in the next paragraph a classical asymptotic result for the sample mean and then we will describe what happens for the sample maxima.

Most statistical inference methods are concerned with the center of a probability distribution (mean) and the deviation from it (variance). One reason for the popularity of such methods is that the assumed distribution is Gaussian, which is uniquely characterized by its mean and its variance. This assumption is justified by a famous theorem in probability, the Central Limit Theorem (CLT), which states the distribution of the sample mean is approximately Gaussian for large samples (whenever the variance is assumed to be finite). The advantage of assuming a Gaussian distribution for the sample mean is that confidence intervals can be then easily derived. Unfortunately, the sample maximum rather than the sample mean has been the primary variable of interest in many recent lichenometry studies (Solomina et al., 2007; McCarroll, 1994). Hence, the CLT can not be applied to approximate the distribution of the maximum (see Fig. 1).

As an alternative, we use a less-known but nevertheless similar result to the CLT. The difference is that this asymptotic result is specially tailored for maxima, and it can be summarized in the following way. The distribution of the maximum in large random samples is known and follows a specific distribution called the Generalized Extreme Value distribution (GEV). The GEV cumulative distribution function is equal to

$$G(x; \mu; \sigma; \xi) = \begin{cases} \exp\left\{-\left[1 + \xi \frac{x - \mu}{\sigma}\right]_+^{-1/\xi}\right\}, & \text{when } \xi \neq 0 \text{ and } a_+ = \max(0, a) \\ \exp\left\{-\exp\left(-\frac{x - \mu}{\sigma}\right)\right\}, & \text{when } \xi = 0 \end{cases} \quad (1)$$

where  $\mu$ ,  $\sigma > 0$  and  $\xi$  are the location, scale, and shape parameters, respectively. The shape parameter drives the tail behavior of  $G$ .

For example, the classical Gumbel distribution corresponds to  $\xi = 0$ . If  $\xi > 0$  (Fréchet type), then the distribution is heavy tailed, i.e. the right tail of the distribution decreases at a much slower rate than a Gaussian tail, and higher moments like the variance may not be finite. If  $\xi < 0$  (Weibull type), it has a bounded upper tail. For example, the solid lines in Figure 1 correspond to GEV densities with different shape parameters. Hence, the sign and value of the shape parameter is very important when modeling the behavior of the maximum. At this juncture, we stress that the distribution of maximum diameters is not chosen arbitrarily; rather the GEV is the unique distribution *derived theoretically* for modeling the maxima. It is important to note that the GEV fits maxima not only from a Gaussian sample, but also from *all* classical continuous distributions (e.g., exponential, uniform, Cauchy, etc.). We refer to the book by Embrechts et al. (1997) for an in-depth discussion about the conditions of this convergence. Hence, this methodology is very general and independent of specific numerical values. Note that the case  $\xi = 0$  in Equation (1) can be viewed as the limiting result:  $\lim_{\xi \rightarrow 0} G(x; \mu, \sigma, \xi) = G(x; \mu, \sigma, 0)$  as  $\xi$  goes to zero. Hence, in the rest of this paper, we will only present our mathematical derivation for the case  $\xi \neq 0$ ; the case  $\xi = 0$  can be obtained by letting  $\xi$  go to zero.

In this paper, we denote the maxima as

$$X_{i,j} = \max_{k=1, \dots, n_{i,j}} \{X_{i,j,k}\}, \text{ for } i = 1, \dots, m, j = 1, \dots, n_i \\ \text{with } n_{i,j} = \sum_k n_{i,j,k} \text{ and } n_i = \sum_j n_{i,j},$$

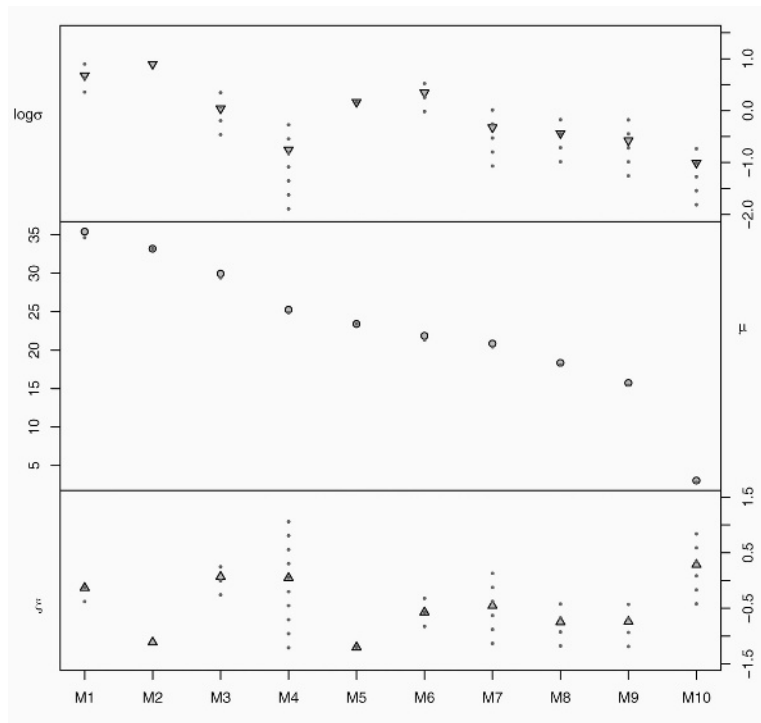
where  $X_{i,j,k}$  are the  $n_{i,j,k}$  lichen diameters on the  $j$ th block from the  $i$ th moraine, and  $m$  is the number of moraines. We suppose that the lichen diameters on a block are independent and follow the same distribution. Between blocks, the lichens are assumed to be independent but they do not have the same distribution. Although we do not necessarily know the value of  $n_{i,j,k}$  and  $X_{i,j,k}$ , we assume that the maximum  $X_{i,j}$  comes from a sample of lichen diameters whose size is relatively large.

From the aforementioned extreme value theory, i.e. Equation (1), it is logical to assume that the probability density function of the maximum lichen diameter  $X_{i,j}$  follows a GEV density,

$$g(x; \mu_i, \sigma_i, \xi_i) = \frac{1}{\sigma_i} \left[ 1 + \xi_i \frac{x - \mu_i}{\sigma_i} \right]_+^{-1/\xi_i - 1} \\ \times \exp\left\{-\left[1 + \xi_i \frac{x - \mu_i}{\sigma_i}\right]_+^{-1/\xi_i}\right\}, \quad (2)$$

for all  $i = 1, \dots, m$  and  $j = 1, \dots, n_i$ . Compared to Equation (1) in which the parameters  $\mu$ ,  $\sigma$ , and  $\xi$  were fixed, Equation (2) shows that the parameters describing lichen diameter distributions can now vary in time through the index of the moraine  $i$  (we suppose that the moraine index  $i$  is ordered in time,  $i = 1$  representing the youngest moraine and  $i = m$  corresponding to the oldest one). The reason for such a choice stems from the very nature of lichenometry, which studies the temporal evolution of large lichen diameter properties from moraine to moraine. Consequently, the behavior of  $\mu_i$ ,  $\sigma_i$  and/or  $\xi_i$  should capture the temporal variation of the lichens. For our case study, which will be described later in this paper, Figure 2 displays the variation of the location, log-scale, and shape parameters in function of the moraines' order. This graph clearly shows that the location parameter varies linearly in function of the moraines' ages, while the scale and shape parameters behave more randomly. In the case study, we will describe how such information can be taken into account in a global statistical model.





**FIGURE 2.** Exploratory study of the maximum lichen diameters measured on ten moraines of Charquini south glacier in Bolivia. The upper, middle, and lower panels correspond to the log-scale, location, and shape GEV parameters [see Equation (2)] in function of the moraine order, respectively, called  $\mu$ ,  $\log \sigma$ , and  $\zeta$ . The moraine M1 represents the youngest one and moraine M10 the oldest. The vertical dotted lines show the 95% confidence intervals. Note that these intervals are very small for the parameter 1a and that they are missing for moraines M2 and M5 due to numerical instabilities.

### DISTRIBUTION OF LICHEN AGES

In lichenometry, there are two categories of dated surfaces. The first one is characterized by the presence of a confidence interval. For example a date computed with a  $^{14}\text{C}$  carbon dating procedure is classically associated with lower and upper bounds. For this category, we can assume that  $T_i$  follows a Gaussian variable with mean  $\alpha_i$  and standard deviation  $\beta_i$ ; the latter parameter being inferred by the  $^{14}\text{C}$  carbon dating procedure. In this case, the density of  $T_i$  is given by

$$f(t_i) = \frac{1}{\sqrt{2\pi\beta_i^2}} \exp\left[-\frac{1}{2\beta_i^2}(t_i - \alpha_i)^2\right]. \quad (3)$$

The second category corresponds to the situation in which the estimated moraine date is not associated with any error analysis, for example a date inferred from a historical map or a human-made structure. In this case, we have two options: either we assume that the estimated date of the moraine is perfect, or we arbitrarily choose to fix a small standard deviation  $\beta$  that characterizes the *a priori* error. For example, even a surface dated from a historical map can be tainted with an error of  $\pm 3$  years with a confidence of 95% (it is difficult to argue that the age of the moraine can be below this error range). Because we prefer to overestimate rather than underestimate the error for our undated surfaces, we opt for the second solution in this study. In this case, the Gaussian density defined by Equation (3) can be used with  $\beta = 3 \text{ years}/1.96 \sim 1.5$ .

### LINKING THE MAXIMUM LICHEN DIAMETERS AND THE MORaine AGES

The last step in our construction is to relate the distribution of the maximum lichen diameters with the moraine ages (the dated and the undated ones). We assume that the joint distribution of  $(X_{i,j}, T_i)$  can be written as the product of a GEV and Gaussian distribution, and the parameters of the two distributions are linked through the

relationship  $\alpha_i = l(\mu_i)$  for some growth function  $l$ . In other words, we suppose that lichen diameters and moraine ages are independent variables (in distribution) but are linked through their parameters. Equations (1) and (2) yield to the joint density function:

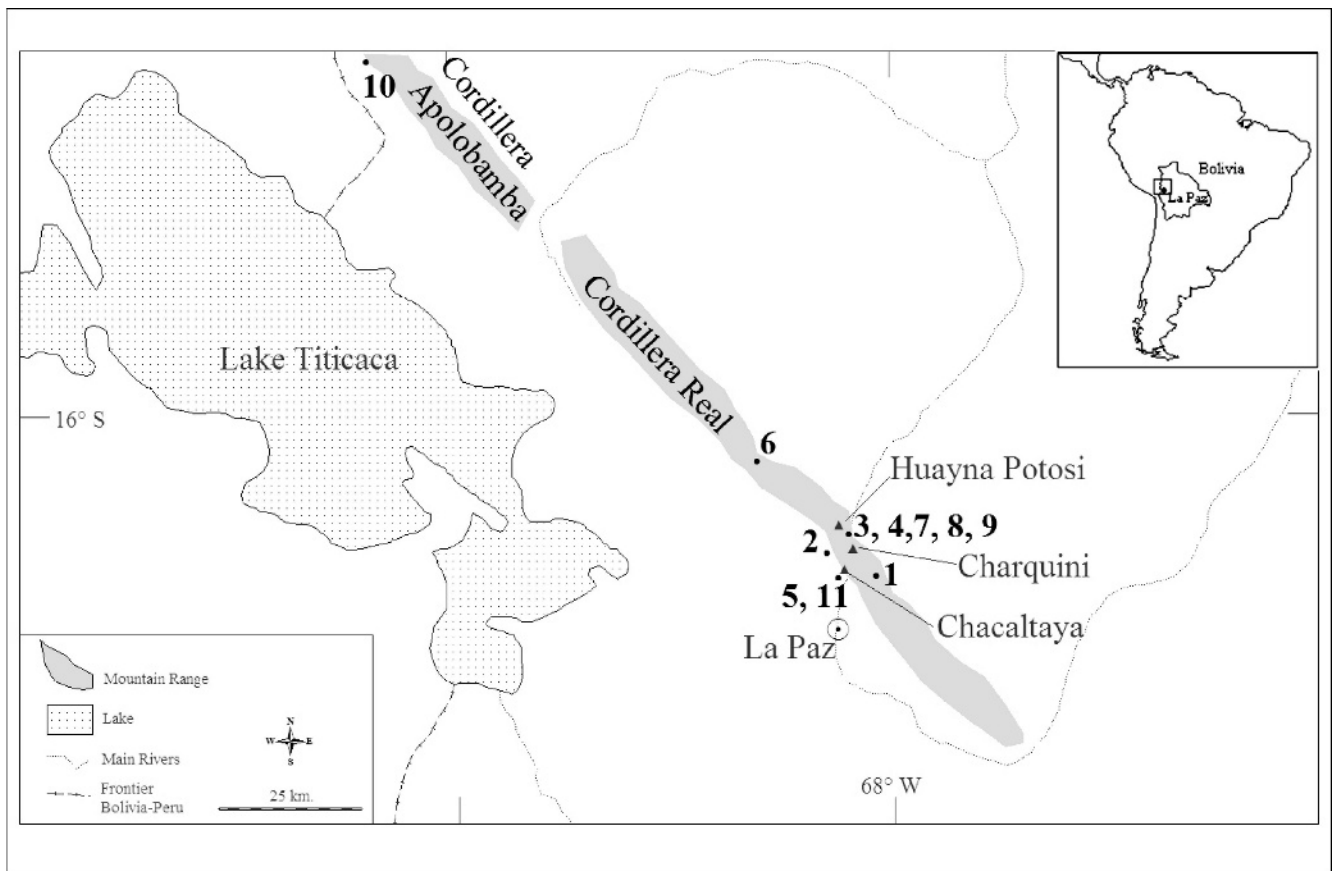
$$f_\theta(x_{i,j}, t_i) = \frac{1}{\sigma_i} \left[ 1 + \zeta_i \left( \frac{x_{i,j} - \mu_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\zeta_i} - 1} \times \exp\left\{ - \left[ 1 + \zeta_i \left( \frac{x_{i,j} - \mu_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\zeta_i}} \right\} \times \frac{1}{\sqrt{2\pi\beta_i^2}} \exp\left[ -\frac{1}{2\beta_i^2}(t_i - \alpha_i)^2 \right], \text{ with } \alpha_i = l(\mu_i), \quad (4)$$

and where  $\theta$  represents the vector of all parameters:  $\theta = [(\mu_1, \mu_2, \dots), (\sigma_1, \sigma_2, \dots), \dots]$ . The above formula is very general and, depending on the case under study, it can be simplified.

### ESTIMATION OF OUR MODEL PARAMETERS

The next stage is to describe a procedure for estimating the unknown parameters represented by the vector  $\theta$  whenever a set of data has been obtained. Suppose that all moraine dates are available. The likelihood function that is defined as the probability of observing our data for a given set of parameters  $\theta$  is equal to

$$l(\theta) = \prod_{i=1}^m \prod_{j=1}^{n_i} f_\theta(x_{i,j}, t_i), \text{ because of the assumed independence among all measurements,} \\ = \prod_{i=1}^m \prod_{j=1}^{n_i} \left( \frac{1}{\sigma_i} \left[ 1 + \zeta_i \left( \frac{x_{i,j} - \mu_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\zeta_i} - 1} \times \exp\left\{ - \left[ 1 + \zeta_i \left( \frac{x_{i,j} - \mu_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\zeta_i}} \right\} \times \frac{1}{\sqrt{2\pi\beta_i^2}} \exp\left[ -\frac{1}{2\beta_i^2}(t_i - \alpha_i)^2 \right] \right), \text{ because of (4).} \quad (5)$$



**FIGURE 3.** The Bolivian region of interest for this study (Cordillera Apolobamba and Cordillera Real). The numbering on this map refers to Table 1.

Under the assumption that all  $t_i$ 's are observable, the optimal parameter vector  $\theta$  is the one that maximizes  $l(\theta)$ . In practice, we only know the  $t_i$ 's for the dated surfaces. To fill in the missing dates, we implement the Expectation-Maximization (EM) algorithm, a recursive process for obtaining maximum likelihood estimates with incomplete data (McLachlan and Thriyambakam, 1997; Dempster et al., 1977). On each iteration of the EM algorithm, there are two steps called the E-step (Expectation) and the M-step (Maximization). This algorithm is closely related to the *ad hoc* approach to estimation with missing data, where the parameters are estimated after filling in initial values for missing values. The latter are then updated by their predicted values using these initial parameters. The parameters are then re-estimated, and so on, proceeding until convergence. The M-step involves only complete-data maximum likelihood estimation, which is computationally simple [see Equation (5)]. The E-step deals with the expectation of the complete-data log likelihood given the observed data (see McLachlan and Thriyambakam, 1997, for more details).

### A Case Study: the Charquini Glacier in Bolivia

#### AREA OF INTEREST

The field work was conducted in the Huayna Potosi–Condoriri massifs (Fig. 3) in the Cordillera Real (16°21'S, 68°07'W) on the eastern part of the Andean chain, about 50 km north of La Paz, Bolivia. To complement our data set, isolated observations were also taken in the northern part of Bolivia in the Cordillera Apolobamba. These areas are dominated by summits reaching 5000 to 6100 m a.s.l. (Huayna Potosi, 6088 m a.s.l.) and comprised of massive batholiths (granite) and metamorphic rock (quartzites).

The climate in this region is defined by the position of the Intertropical Convergence Zone (ITCZ), the oscillation responsible for a marked rainy seasonal variability in the eastern Andean area (Aceituno, 1988; Ribstein et al., 1995; Vuille et al., 1998; Garreaud, 1999). The southern winter (May to September) produces a dry and cold season generated by the northward displacement of the mid- and upper tropospheric westerlies. The southern summer (November to March) is warm and wet. This area within the ITCZ is characterized by low seasonal variations in solar radiation and temperature and by a marked seasonality in precipitation. Precipitation was measured on the Plataforma Zongo, 2 km from Charquini, where the annual average precipitation from 1971 to 2000 was about 835 mm at 4800 m a.s.l. (Caballero et al., 2002). Around 65% of the rain falls from December to February. Monthly average temperature variations do not present a large amplitude. The 0°C isotherm remains above 4900 m throughout the year. A recent study of the snow cover duration near the glacier studied therein showed that snow represents around 26% of precipitation and that the snow cover at 4900 m during the rainy season stays on the ground rarely more than three days, with a modal value of one day and a maximum of six days (Chevallier, 2002). The Equilibrium Line Altitude of glaciers is between 5200 and 5400 m a.s.l. The periglacial environment is relatively spread in altitude, and permafrost may exist locally from 5400 m a.s.l. (Francou et al., 2001; Ramirez et al., 2001).

#### FIELD WORK

Our data set of lichenometric measurements of *Rhizocarpon* sp. (ex *Rhizocarpon geographicum*) was divided into two groups: 10

**TABLE 1**  
**Characteristics of the dated surfaces.**

Sites in Fig. 3	Surface	Lichen numbers	Block numbers	z	Exposition	Age (A.D.)	Type <sup>1</sup>	Reference
1	Rail road	153	41	4650	all	1915	OD	Cobee company
2	Milluni Dam	174	20	4650	E	1911	OD	Cobee company
3	Huayna moraine	218	10	4860	E	1956	AP	Bolivian Military Lib
4	Zongo moraine	200	11	4880	E	1956	AP	Bolivian Military Lib
	Zongo moraine	200	11	4880	E	1963	AP	Bolivian Military Lib
5	Chacaltaya moraine	229	10	4900	S	1948	AP	Ramirez et al. (2001)
	Chacaltaya moraine	234	10	4900	S	1941	P	Ramirez et al. (2001)
6	Janqu Uyu moraine	1638	12	4800	W	1630–1670 <sup>2</sup>	<sup>14</sup> C	Gouze et al. (1986)
7	Zongo road	60	12	4650	all	1928	OD	Cobee company
	Zongo road	60		4400	all	1935	OD	Cobee company
8	Zongo Canal	158		4800	NE	1946	OD	Cobee company
9	Charquini Canal	167	15	4800	NE	1941	OD	Cobee company
10	Suchez houses	451	10	4800	all	1767	OD	Herail et al. (1985)
11	Chacaltaya houses	214	10	4850	all	1935	OD	Cobee company

<sup>1</sup> Type: OD = old documents, AP = aerial photos, and P = photos.

<sup>2</sup> Highest probability for two sigma obtained with Calib 4.4 (Stuiver et al., 1998)

dated surfaces and 10 undated moraines. The latter group constituted the main focus of our investigation, the undated moraines on the Charquini glacier. All measurements on this glacier were made on the external side of lateral moraines on both sides of the valley (see Table 2 for the sampling characteristics). We selected at least 180 blocks per moraine with 90 on either side of the valley. For each block, the single largest lichen was measured. The lichen size ranged from 10 to 50 mm, while the block position with respect to the moraine axis was noted. We were not able to detect any significant spatial variation from block to block.

For the second group, different dating techniques (historical documents, aerial photos, carbon dating, etc.) were employed to derive associated dates. Table 1 summarizes the results of this dating procedure; the fact that the same date could be obtained by different techniques explains the redundancy of a few dates in our table. Geographically, six dated moraines from four glaciers were located in the Huyana Potosi Massif (Charquini, Chacaltaya, Janqu Uyu, and Zongo). The oldest moraine of our sample was carbon dated (<sup>14</sup>C from peat samples; see Table 1 and Stuiver et al., 1998). The known spread was included in our statistical analysis by fixing the dating range for this moraine. Otherwise we assumed that the standard deviation  $\beta$  in Equation (3) of all the other dated surfaces in Table 1 was fixed and equal to 1.5 years. In Table 1, there was also a category of surfaces comprised of a variety of human-made structures such as archeological monuments, rock-wall cut, stonewalls during the construction of a road, artificial dams, and irrigation canals. For these diverse surfaces, the sampling was performed in the following way. In the case of rocky slopes, the single largest lichen was selected inside an area of 1 m<sup>2</sup>. This procedure was repeated on 60 different sampled surfaces. In the case of housing walls, measurements were taken on inside walls as well as outside ones, but no difference was identified between them. Similar measurements were performed on the Milluni dam that was constructed from stones of 30–50 cm length along the a-axis. We also tried to take measurements on walls of old buildings like colonial churches but all these trials were unsuccessful. Finally, all measurements were taken with a flexible, transparent plastic rule with an accuracy of 1 mm. The smallest measured diameter was 2 mm. Anomalous lichen shapes were rejected to reduce the risk of coalescence. At the end of this field campaign, around 6800 measurements were obtained, i.e. 4156 for dated surfaces and 2600 on the Charquini glacier.

## STATISTICAL ANALYSIS OF THE BOLIVIAN DATA

### Exploratory Analysis of the GEV Fit with Respect to the Moraine Order

As a preliminary step before modeling all the elements of the likelihood function defined by Equation (5), it is interesting to have a crude idea of the temporal variation of the three GEV parameters. For each of the 10 moraines found on the Charquini south glacier, the maximum diameter lichen distribution is independently fitted by maximizing the likelihood obtained from Equation (2). Figure 2 displays estimates of the GEV location, shape, and log-scale parameter for each moraine. The vertical dotted lines correspond to the 95% confidence intervals. Each of the three panels is discussed separately in the following paragraphs.

The GEV location parameter  $\mu_i$  in Equation (2) corresponds to the main temporal trend in maximum lichen diameters. For example, a linear trend should fit adequately our Bolivian data (see the middle panel of Fig. 2). In general, the  $\mu_i$ 's play an equivalent role of the growth curve. After a rapid increase during the earlier part of the lichen life caused by competitive stress, the growth curve follows a roughly linear trend, and finally its value converges to a plateau which reflects the slowing down of the lichen growth. These different phases can be parameterized by a mathematical formula depending on the time period under study. For our Bolivian lichen diameters, the middle panel of

**TABLE 2**  
**Characteristics of the lichen sampling for the Charquini south glacier.**

Moraine number	Number of lichen	z	Number of blocks
M1	513	4762	30
M2	368	4765	16
M3	256	4766	16
M4	244	4776	16
M5	189	4800	12
M6	257	4825	30
M7	221	4837	15
M8	191	4842	27
M9	186	4872	16
M10	180	4905	15



TABLE 3

Charquini south glacier: Comparison among different GEV models defined by Equation 5. Four models for the log-scale shape parameter  $\log \sigma_i$  are compared in this table. The location  $\mu_i$  parameter is represented as a linear function, and the shape parameter  $\xi_i$  is assumed to be a constant.

$\mu_i$	$\log \sigma_i$	$\xi_i$	log-likelihood: nb of AIC:		
			$\log l(\theta)$	parameters: $k$	$-2 \log l(\theta) + 2k$
$a_\mu + b_\mu i$	$\log \sigma_i$	$\xi$	-741	$2 + 21 + 1 = 24$	1530
$a_\mu + b_\mu i$	$a_\sigma + b_\sigma i + c_\sigma i^2$	$\xi$	-826	$2 + 3 + 1 = 6$	1664
$a_\mu + b_\mu i$	$a_\sigma + b_\sigma i$	$\xi$	-1084	$2 + 2 + 1 = 5$	2178

Figure 2 clearly shows that our data span the linear period of the lichen growth.

In contrast to the location parameter, the GEV shape parameter  $\xi_i$  represents the large-scale behavior of lichen distributions. For our particular case study in Bolivia, there is no reason to think the moraines had been perturbed by large-scale changes in the last few centuries. In addition, the lower panel of Figure 2 clearly shows two features. First, the shape parameter does not display a clear temporal pattern, and second, there is a large variability among our individual estimated shape parameter values. This last issue is due to the statistical difficulty of estimating  $\xi_i$ 's for small samples, i.e. estimating separately  $\xi_i$  for each moraine  $i$ . For moraines M2 and M5, it is not even possible to compute confidence intervals due to numerical instabilities in the likelihood estimation procedure of  $\xi_i$ . These two reasons justify our main model choice. We assume that  $\xi_i$  is a constant for all moraines found on the Charquini south glacier. Of course, if there is geophysical or climate evidence that an abrupt temporal change had occurred for a particular glacier, then this information should be integrated in our model.

The scale parameter  $\sigma_i$  has to obey a mathematical constraint; it must be positive. For this reason,  $\log \sigma_i$  is classically parameterized (instead of  $\sigma_i$ ). As seen in Figure 2, it may be possible that  $\log \sigma_i$  follows a particular trend, but it is very difficult to visually decide if such a possible trend is significant. To solve this issue, we need to introduce a few concepts about model selection.

Model Selection

From Figure 2, we learned that a linear parameterization is reasonable for modeling the location parameter. Hence, our general likelihood function defined by Equation (5) can be simplified as

$$\begin{aligned}
 l(\theta) = & \prod_{i=1}^m \prod_{j=1}^{n_i} \left( \frac{1}{\sigma_i} \left[ 1 + \xi \left( \frac{x_{i,j} - \alpha_\mu - b_\mu \alpha_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\xi} - 1} \right. \\
 & \times \exp \left\{ - \left[ 1 + \xi \left( \frac{x_{i,j} - \alpha_\mu - b_\mu \alpha_i}{\sigma_i} \right) \right]_+^{-\frac{1}{\xi}} \right\} \\
 & \times \frac{1}{\sqrt{2\pi\beta_i^2}} \exp \left[ - \frac{1}{2\beta_i^2} (t_i - \alpha_i)^2 \right],
 \end{aligned} \tag{6}$$

where the  $\beta_i$ 's are given by Calib 4.4 (Stuiver et al., 1998) from  $^{14}\text{C}$  carbon dates.

To completely define our likelihood function, we still need to propose a parameterization for the log-scale parameter  $\log \sigma_i$  in Equation (6). In Table 3, we propose and compare three different models for the log-scale parameter. The simplest one corresponds to a linear model for  $\log \sigma$  and for the most complex, we suppose a different  $\log \sigma$  for each moraine. As an intermediate step, we also

TABLE 4

Estimated ages, Charquini south glacier.

Moraine number	Estimated age	95% confidence interval
M1	1720	24
M2	1736	23
M3	1752	22
M4	1784	20
M5	1800	19
M6	1816	18
M7	1832	17
M8	1848	16
M9	1864	15
M10	1896	13

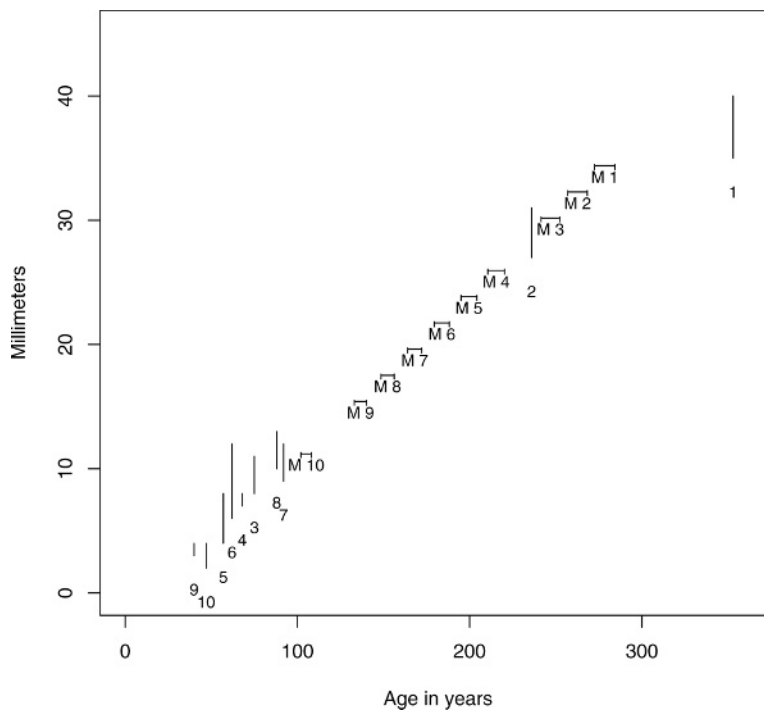
investigate a quadratic trend. An important issue related to the choice of the appropriate model with respect to our data is the problem of over-parameterization, fitting the data very well but with a large number of parameters. For example, 10 data points can be perfectly modeled by 10 parameters but the predictive value of such an over-parameterized model is null. To compare the models described by the first two columns of Table 3, we compute the Akaike Information Criterion (AIC) (George, 2000) that is defined by  $AIC = -2 \log l(\theta) + 2k$ , where  $k$  is the number of parameters in our model (see Table 3). Heuristically, one can view the first term,  $-2 \log l(\theta)$ , as a measure of lack of fit, while the second term,  $2k$ , can be interpreted as a ‘‘penalty’’ for increasing the number or the size of the model (the penalty enforces parsimony in the number of parameters). According to Table 3 and the AIC, the model with a different  $\log \sigma_i$  for each moraine is the best one among our three investigated models.

Moraine Ages Estimates

Table 4 provides the estimated ages and their 95% associated confidence intervals for model 1 in Table 3. As a supplement of Table 4, Figure 4 displays the estimated ages in function of the estimated location parameter  $\mu$  (see the horizontal lines). The width of these horizontal segments shows the 95% confidence interval length of the estimated ages. The vertical lines represent the range of the dated surfaces. Again, we would like to stress that these 95% intervals are not comparable to the ones obtained with past lichenometric approaches. These past methods were based on the false assumption of a Gaussian distribution despite the fact that probability theory and practice (see Fig. 1) dictate that the shape of the GEV distribution is more appropriate.

Conclusions and Future Work

Besides the small size of the computed errors in Figure 4, the main improvement over past methods is that the proposed model is more general and more theoretically sound. Still, a number of important statistical questions remains in lichenometry. The model we proposed here only works for block maxima. However, it may be more statistically robust to sample not only the largest size per block but the second or third largest lichens. Determining a mathematical criterion for choosing the most appropriate sampling scheme remains a fundamental question in lichenometry. We believe that extreme value theory can be a valuable tool to perform such a task. Another interesting challenge is determining how to pool data coming from glaciers in the same vicinity. For example, one can imagine that a large quantity of lichen diameters are measured on moraines located on a glacier near the Charquini glacier. If these two glaciers are geographically close, they should share a range of



**FIGURE 4.** A new type of growth curve plot. The horizontal lines correspond to the estimated location parameters  $\mu$  of the GEV distribution [see Equations (1) and (2)] for each of the 10 moraines. The width of these horizontal segments shows the 95% confidence interval length of the estimated ages. The vertical lines represent the range of the dated surfaces.

similar characteristics and this could improve the error analysis. In the near future, we plan to investigate these two issues.

Due to the theoretical origin of the GEV distribution, it is difficult to dispute our modeling of maximum diameters; however, it is important to recognize that our proposed global model and estimation strategy are not unique. One may prefer to construct a different link between dates and diameters. For example, a Bayesian approach could also be implemented or a hidden Markov model could be used to model this relationship. Our objective in this paper was to show that a “simple” parametric model could perform reasonably well for fitting our data set, but other levels of complexity (if some of our assumptions are not true or more information is available) could be added and lead to a slightly different error analysis.

Finally, we are currently working on making available to the lichenometry community a user-friendly program on the web that will allow the practitioner to test our approach with their own data.

### Acknowledgments

This research was supported by the IRD, the CNRS program ECLIPSE, the NFS grant ATM-0327936, the NSF grant VIGRE, and the E2C2 European project. In addition, we would like to thank the anonymous reviewers for their comments, which greatly improved the article.

### References Cited

Aceituno, P., 1988: On the functioning of the southern oscillation the south American sector, part I. Surface climate. *Monthly Weather Review*, 116: 505–524.

Benedict, J., 1967: Recent glacial history of an alpine area in the Colorado Front Range, USA. I. Establishing a lichen growth curve. *Journal of Glaciology*, 6: 817–832.

Benedict, J., 1990: Lichen mortality due to late-lying snow, results of a transplant study. *Arctic and Alpine Research*, 22: 81–99.

Benedict, J., 1991: Experiments on lichen growth ii. Effects of a seasonal snow cover. *Arctic and Alpine Research*, 23: 189–199.

Beschel, R., 1961: Dating rock surfaces by lichen growth and its application to glaciology and physiography (lichenometry). In Raasch, G. O. (ed.), *Geology of the Arctic*. Volume 2. Toronto: University of Toronto Press, 1044–1062.

Bull, W., and Brandon, M., 1998: Lichen dating of earthquake-generated regional rockfall events, Southern Alps, New Zealand. *Geological Society of America Bulletin*, 101: 60–84.

Caballero, Y., Jomelli, V., and Chevallier, P., 2002: Hydrological characteristics of slope deposits in high tropical mountains (Cordillera Real, Bolivia). *Catena*, 47(2): 101–116.

Chevallier, P., 2002: Dynamique de la couverture neigeuse dans les Andes tropicales. Programme de Recherche (2001–2002) PNRH CNRS, unpublished Rapport, 17 pp.

Coles, S. G., 2001: *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. London: Springer-Verlag London Ltd.

Dempster, A., Laird, N., and Rubin, D., 1977: Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39B: 1–38.

Embrechts, P., Klüppelberg, C., and Mikosch, T., 1997: *Modelling extremal events for insurance and finance*. Volume 33, Applications of Mathematics. Berlin: Springer-Verlag.

Francou, B., Mehaute, N. L., and Jomelli, V., 2001: Factors controlling spacing distances of sorted stripes in a low-latitude, alpine environment (Cordillera Real, Bolivia, 16S). *Permafrost and Periglacial Processes*, 12: 367–377.

Gaines, S., and Denny, M. W., 1993: The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. *Ecology*, 14(6): 1677–1692.

Garreaud, R., 1999: Multi-scale analysis of the summertime precipitation over the central Andes. *Monthly Weather Review*, 12: 3157–3171.

George, E., 2000: The variable selection problem. *Journal of the American Statistical Association*, 95: 1304–1308.

Gnedenko, R., 1943: Sur la distribution limite du terme maximum d’une série aléatoire. *Annals of Mathematics*, 44: 423–453.

Gouze, P., Argollo, J., Saliège, J., and Servant, M., 1986: Interpretation paléoclimatique des oscillations des glaciers au cours des 20 derniers millénaires dans les régions tropicales; exemple des Andes boliviennes. *Comptes rendus de l’Académie des Sciences Paris*, 303(Série II): 219–224.

- Heraïl, G., Fornari, M., Viscarra, G., Ruiz, J., Pozzo, L., and Dumont, J.-F., 1985: Les placers d'or de Bolivie, milieux de formation et structure géologique. Gisement alluviaux d'or. Actes du symposium international sur les gisements alluviaux d'or. La Paz, Bolivia, 3–5 Juin 1991, Orstom ed., 2115–2143.
- Karlen, W., and Black, J., 2002: Estimates of lichen growth-rate in northern Sweden. *Geografiska Annaler*, 84A: 225–232.
- Katz, R., Parlange, M., and Naveau, P., 2002: Extremes in hydrology. *Advances in Water Resources*, 25: 1287–1304.
- Kharin, V., and Zwiers, F., 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean gcm. *Journal of Climate*, 13: 3760–3788.
- Leadbetter, M., Lindgren, G., and Rootzén, H., 1983: *Extremes and related properties of random sequences and processes*. New York: Springer Verlag.
- McCarroll, D., 1994: A new approach to lichenometry, dating single-age and diachronous surfaces. *The Holocene*, 4: 383–396.
- McLachlan, G., and Thriyambakam, K., 1997: *The EM algorithm and extensions*. New York: Wiley.
- Naveau, P., and Moncrieff, M., 2003: A statistical formulation of convective mass fluxes. *Quarterly Journal of the Royal Meteorology Society*, 129: 2217–2233.
- Naveau, P., Nogaj, M., Ammann, C., Yiou, P., Cooley, D., and Jomelli, V., 2005: Statistical analysis of climate extremes. *Comptes rendus Geosciences de l'Academie des Sciences*, 337: 1013–1022.
- O'Neal, M., and Schoenenberger, K., 2003: A *Rhizocarpon geographicum* growth curve for the Cascade Range of Washington and northern Oregon, USA. *Quaternary Research*, 60: 233–241.
- Pentecost, A., 1979: Aspect and slope preferences in a saxicolous lichen community. *Lichenologist*, 11: 81–83.
- Reiss, T., 1997: *Statistical analysis of extreme values*. Basel: Birkhauser.
- Ribstein, P., Tiriau, E., Francou, B., and Saravia, R., 1995: Tropical climate and glacier hydrology, a case study in Bolivia. *Journal of Hydrology*, 165: 221–234.
- Rodbell, D., 1992: Lichenometric and radiocarbon dating of Holocene glaciation, Cordillera Blanca, Peru. *The Holocene*, 2: 19–29.
- Smith, R., 1992: Extreme value theory. In Ledermann, W. (ed.), *Handbook of applicable mathematics*. Volume 7. Chichester: John Wiley, 437–471.
- Solomina, O., Jomelli, V., Kaser, G., Ames, A., Berger, B., and Pouyaud, B., 2007: Little Ice Age moraines in the Cordillera Blanca, lichenometric data replication. *Global and Planetary Change*. in press.
- Stuiver, M., Reimer, P., Bard, E., Beck, J., Burr, G., Hughen, K., Kromer, B., McCormac, F., Plicht, J., and Spurk, M., 1998: Intcal98, radio carbon age calibration 24,000-o cal bp. *Radiocarbon*, 40: 1041–1083.
- Vuille, M., Hardley, D., Braun, C., Keimig, F., and Bradley, R., 1998: Atmospheric circulation anomalies associated with 1996–1997 summer precipitation events on Sajama ice cap, Bolivia. *Journal of Geophysical Research*, 103: 11191–11204.

Ms accepted September 2006