Extreme value analysis of wildfires in Canadian boreal forest ecosystems

Yueyang Jiang and Qianlai Zhuang

Abstract: Large fires are a major disturbance in Canadian forests and exert significant effects on both the climate system and ecosystems. During the last century, extremely large fires accounted for the majority of Canadian burned area. By making an instaneous change over a vast area of ecosystems, extreme fires often have significant social, economic, and ecological consequences. Since extreme values of fire size always situate in the upper tail of a cumulative probability distribution, the mean and variance alone are not sufficient to fully characterize those extreme events. To characterize the large fire behaviors in the upper tail, the authors in this study applied three extreme value distribution functions: (*i*) the generalized extreme value (GEV) distribution, (*ii*) the generalized Pareto distribution (GPD), and (*iii*) the GEV distribution with a Poisson point process (PP) representation to fit the Canadian historical fire data of the period 1959–2010. The analysis was conducted with the whole data set and different portions of the data set according to ignition sources (lightning-caused or human-caused) and ecozone classification. It is found that (*i*) all three extreme statistical models perform well to characterize extreme fire events, but the GPD and PP models need extra care to fit the nonstationary fire data, (*ii*) anthropogenic and natural extreme fires have significantly different extreme statistics, and (*iii*) fires in different ecozones exhibit very different characteristics in the view of statistics. Further, estimated fire return levels are comparable with observations in terms of the magnitude and frequency of an extreme event. These statistics of extreme values provide valuable information for future quantification of large fire risks and forest management in the region.

Résumé : Les grands feux constituent une perturbation majeure dans les forêts canadiennes et ont des effets importants tant sur le système climatique que sur les écosystèmes. Au cours du dernier siècle, la majorité des superficies brûlées au Canada l'ont été lors de feux extrêmement importants. En provoquant des changements instantanés sur de vastes superficies des écosystèmes, les feux extrêmes ont souvent d'importantes conséquences sociales, économiques et écologiques. Étant donné que les valeurs extrêmes de la dimension des feux se situent toujours à l'extrémité supérieure d'une distribution de probabilité cumulative, la moyenne et la variance seules ne suffisent pas pour caractériser pleinement ces événements extrêmes. Pour caractériser le comportement des grands feux dans l'extrémité supérieure de la distribution, les auteurs de cette étude ont appliqué trois fonctions de distribution de valeurs extrêmes : (*i*) la distribution généralisée de valeurs extrêmes (DGVE), (*ii*) la distribution Pareto généralisée (DPG) et (*iii*) la DGVE avec une représentation par le processus ponctuel de Poisson (PP) pour décrire les données canadiennes de l'historique des feux durant la période 1959 à 2010. L'analyse a été réalisée avec le jeu de données au complet et avec différentes portions selon la source d'allumage (feu causé par la foudre ou par l'homme) et la classification de l'écozone. On constate que (*i*) les trois modèles de statistiques extrêmes sont capables de bien caractériser les épisodes de feu extrême mais les modèles DPG et PP nécessitent une attention particulière pour ajuster les données de feu évolutif, (*ii*) les feux extrêmes d'origine humaine et naturelle ont des statistiques extrêmes significativement différentes et (*iii*) compte tenu des statistiques, les caractéristiques des feux sont très différentes selon l'écozone. De plus, les niveaux estimés de retour du feu sont comparables aux observations en termes d'ampleur et de fréquence d'un événement extrême. Ces statistiques de valeurs extrêmes fournissent une information précieuse pour la quantification future des risques de grand feu et pour l'aménagement forestier dans la région.

[Traduit par la Rédaction]

Introduction

Large fires ($\geq$2 km^2) are the most important disturbances and account for approximately 97% of the total burned area of Canadian forests, although they only represent 3% of the total fire occurrences (Stocks et al. 2002). Previous studies have shown that burned areas significantly increased in the last four decades (e.g., Podur et al. 2002; Gillett et al. 2004; Kasischke and Turetsky 2006; Xiao and Zhuang 2007) and this trend will continue in the 21st century under the future warming condition (e.g., ACIA 2004; Flannigan et al. 2005; Zhuang et al. 2006). Extreme fire events often cause high economic and human costs (Alvarado et al. 1998) and landscape heterogeneity resulting from extreme fires can contribute to habitat diversity of boreal landscapes (Burton et al. 2008) as well. An adequate characterization of extreme fires

Received 9 June 2011. Accepted 21 June 2011. Published at www.nrcresearchpress.com/cjfr on 29 August 2011.

Y. Jiang. Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907, USA.
Q. Zhuang. Department of Earth and Atmospheric Sciences and Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA.

Corresponding author: Yueyang Jiang (e-mail: jiang5@purdue.edu).

is crucial to fire prediction and management (Flannigan et al. 1998; Stocks et al. 1998, 2001; Wotton et al. 2003; Beverly and Martell 2005).

Analyses using the mean and variance of fire sizes or frequency (e.g., Kushla and Ripple 1997) are not sufficient to investigate the uncertainty of fire distribution because a single extreme event might dramatically disrupt the central tendency of the fire occurrence distribution (Alvarado et al. 1998). To date, very few analyses have focused on the upper tail of fire size distribution, which is occupied by extreme fire events. Previous studies (e.g., Malamud et al. 2005; Jiang et al. 2009) suggested that heavy-tailed distributions are suitable for characterizing large fire behaviors. However, because extreme fire events are rare, the shape of the tail is often not well characterized (Malamud et al. 2005). Several studies (e.g., Cumming 2001; Reed and McKelvey 2002; Schoenberg et al. 2003) provided some methods to estimate the upper quantiles (extreme values). For example, estimates for upper quantiles might be obtained from quantile regression (Koenker 2005) without specifying a parametric distribution. Back-transformation methods like generalized additive models for location, scale, and shape are also available for estimating upper quantiles (Rigby and Stasinopoulos 2005).

In this study, we applied the extreme value theory (EVT) to analyze extremely large fires with parametric models that could be further used to predict fire behaviors (e.g., fire size and frequency) with a consideration of spatial and temporal covariates (e.g., climate and fuel load) that affect fire behaviors (Cumming 2001). The EVT has been used to model ecological disturbances (Alvarado et al. 1998; Katz et al. 2005). Here, we used the generalized extreme value (GEV) distribution, the generalized Pareto distribution (GPD), and the GEV distribution via the Poisson point process (PP) approach (Coles 2001) to fit the time series of Canadian fire size data set from 1959 to 2010. The EVT focuses on the upper tail of the fire size distribution that is occupied by very large fires. For the large but not necessarily extremely large fire events (e.g., 0.90–0.95 quantiles), other approaches (e.g., the power–law frequency–area model: Malamud et al. 2005; Jiang et al. 2009) have proven to be sufficient to model these large fire events.

Since characterizing extremely large fires induced from various ignition sources is important for fire prediction (Nash and Johnson 1996; Dey and Guyette 2000; Guyette et al. 2002; Krawchuk et al. 2006; Jiang et al. 2009; Krawchuk and Cumming 2009), we separately modeled the anthropogenic and natural extreme fires and compared their statistics of the extreme values. In addition, since fire behaviors could be influenced by different factors, such as weather conditions, vegetation cover. and topography, we conducted two separate extreme value analyses on extreme fires that occurred in the two largest Canadian ecozones: the Boreal Shield ecozone and the Taiga Shield ecozone. These two ecozones have different vegetation cover, topography, density of rivers and lakes, and intensity of fire suppressions, all of which could exert significant effects on fire behavior (Cumming 2001). The majority of large fires also occurred in these two ecozones in the period 1959–2010.

Next, based on the developed extreme statistical models, the *m*-year fire return level defined as the burned area that is expected to exceed a threshold once in a region, on average, during the *m*-year period was estimated. The implication of the characterization of these extreme fires and the estimated fire return levels to fire prediction and management is discussed.

Data preparation

The Canadian National Fire Database (CNFDB) of the Canadian Forest Service was used to investigate the extreme fire statistics. The CNFDB is a collection of forest fire locations and fire perimeters data provided by Canadian fire management agencies including provinces, territories, and Parks Canada (Burton et al. 2008; Canadian Forest Service 2010). Around 300 000 fires are recorded in the CNFDB for the period from 1959 to 2010 including the Large Fire Database (1959–1999) (Stocks et al. 2002). Limitations of the CNFDB include (*i*) fire locations are approximated, (*ii*) data completeness and quality vary depending on mapping techniques, source agencies, and fire years, and (*iii*) data collected in more recent years may be more reliable. There is a large interannual variability in both the number of fires and area burned in Canada (Fig. 1).

We processed the fire data by summing up all fires started on the same day to produce the daily time series for our analysis. This overcomes some geographic biases. The fire start date was used as the signature of the daily fire event, but it does not mean that all of the area burned happened in a single day. Instead, the ignitions that occurred on the same day result in a certain fire size, even an extremely large one. The justification for using the daily sums approach is still debatable; we thus also processed the data set into monthly time series by summing up fire sizes in the same month. After processing, there are in total 18 993 daily-organized events and 624 monthly-organized events for the EVT analysis at two different temporal scales.

Methods

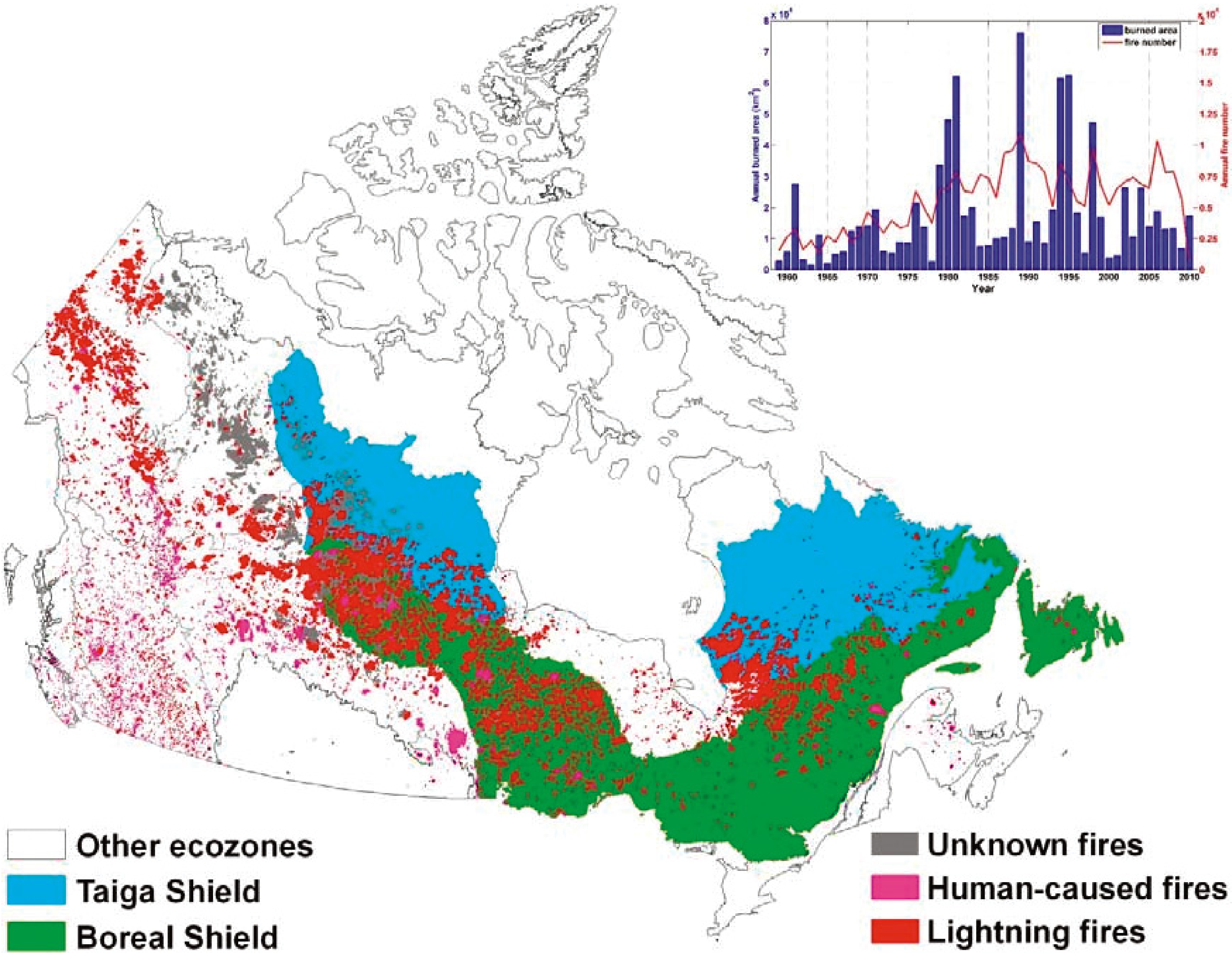
The statistical theory of extreme events

In this study, three extreme value models were used to model the large fires as extreme events: (*i*) the GEV distribution, (*ii*) the GPD, and (*iii*) the GEV distribution with a PP representation. Details about these three models are presented in Coles (2001).

The cumulative density function (CDF) of the GEV distribution is

$$
[1] \qquad G(z) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & 1+\xi\left(\frac{z-\mu}{\sigma}\right) > 0, \xi \neq 0 \\ \exp\left[-\exp\left(-\frac{z-\mu}{\sigma}\right)\right] & \xi = 0 \end{cases}
$$

Fig. 1. Ecozone boundaries and geographical distribution of fires ($\geq$2 km^2) across Canada from 1959 to 2010. The greyscale polygons (colored online) represent the burned areas. The chart in the upper right corner shows statistics extracted from the Canadian National Fire Database (CNFDB). The burned areas may be overlapped by one another.



where μ represents a location, σ is a scale, and ξ is a shape parameter. The shape of the GEV distribution assumes three possible types based on the value of ξ: (*i*) $\xi = 0$, a light-tailed (or Gumbel) distribution, (*ii*) $\xi > 0$, a heavy-tailed (or Fréchet) distribution, and (*iii*) $\xi < 0$, a bounded (or Weibull) distribution.

The GPD is recommended (e.g., Davison and Smith 1990) to describe excesses of a high threshold u, $Y = X–u$ of which the CDF is

$$[2] \qquad H(y) = \begin{cases} 1 - \left(1 + \dfrac{\xi y}{\sigma^*}\right)^{-1/\xi} & 1 + \dfrac{\xi y}{\sigma^*} > 0, \xi \neq 0 \\ 1 - \exp\left(-\dfrac{y}{\sigma^*}\right) & \xi = 0 \end{cases}$$

where $y > 0$, scale $\sigma^* > 0$, and ξ is the shape parameter. Similar to the GEV distribution, the shape parameter determines three possible types: (*i*) $\xi = 0$, a light-tailed (or exponential) distribution, (*ii*) $\xi > 0$, a heavy-tailed (or Pareto) distribution, and (*iii*) $\xi < 0$, a bounded (or beta) distribution.

The PP representation provides a formal theoretical justification for the peak over threshold (POT) method of fitting the GEV distribution (Katz et al. 2005). As mentioned in Katz et al. (2005), the PP representation is a two-dimensional nonhomogeneous Poisson process that combines the Poisson process (parameter λ) for the times of exceedance of the high threshold and the GPD (parameters σ^* and ξ) for the excesses over the threshold (Leadbetter et al. 1983; Smith 1989; Davison and Smith 1990). The relationship between parameters λ and σ^* and the GEV distribution parameters μ, σ, and ξ is described as

$$[3] \qquad \ln\lambda = -\frac{1}{\xi}\ln\left[1 + \xi\left(\frac{u-\mu}{\sigma}\right)\right]$$
$$\sigma^* = \sigma + \xi(u - \mu)$$

The two distributions (GPD and PP) have an identical shape parameter ξ.

In practice, there are two issues in fitting the GPD and PP models to extreme fire events. The first issue is the dependence of extremes. The classical EVT only applies to independently and identically distributed data. However, large fire events may show clusters of observations in the tail and the potential correlation among threshold excesses could lead to unrealistically tight confidence intervals. In fact, the CNFDB is best described as a nonstationary time series with unknown dependence structure. Unfortunately, there is no general theory to deal with this nonstationary process. Nevertheless, as demonstrated in Coles (2001), the standard extreme value models could be used as basic templates that can be enhanced by statistical modeling.

To deal with the problem of dependent threshold excesses, the most widely adopted method is declustering. Following Coles (2001), we declustered the data set to obtain a set of exceedances that are approximately independent. First, we specified a threshold u and considered a cluster to be active until r consecutive values fall below u. Here, we defined u equal to the 0.99 quantile of fire data. As noted in Coles (2001), there is a trade-off for the choice of r between bias and variance: too small a value will lead to the problem of independence being unrealistic for nearby clusters; too large a value will lead to a concatenation of clusters that could reasonably have been considered as independent and therefore to a loss of valuable data. To obtain the estimated optimal r (run length), we estimated the extremal index following Ferro and Segers (2003). The extremal index is an indicator of how much clustering of exceedances of a threshold occurs in the limit of the distribution. If the extremal index is equal to or larger than 1, then it suggests no dependence at extreme levels; otherwise, if it is less than 1, there is some dependency in the limit.

The second issue is the choice of threshold for the POT method. Both the GPD and PP approaches require choosing a high threshold. However, the threshold selection is still an unsolved problem, but some diagnostic tools exist (Coles 2001) that assess features of the model fit for a range of potential thresholds. The threshold was determined through a trade-off between the reliability of the asymptotic approximation versus the variance of estimators. On the one hand, the threshold should be high enough to guarantee that the distribution of exceedances is in the domain of attraction of the generalized Pareto family; on the other hand, the selected threshold ensures that the number of exceedances is not too few to guarantee accurate estimates of parameters. A sensitivity test was conducted to investigate how parameter estimates are sensitive to a range of thresholds.

The return level Z_p, associated with the return period $1/p$, is equivalent to the $(1 - p)$th quantile of the corresponding CDF. For the GEV distribution, it can be calculated by inverting eq. 1:

$$[4] \qquad z_p = \begin{cases} \mu - \dfrac{\sigma}{\xi}\left\{1 - [-\log(1-p)]^{-\xi}\right\} & \text{for } \xi \neq 0 \\ \mu - \sigma \log[-\log(1-p)] & \text{for } \xi = 0 \end{cases}$$

It states that the annual maximum burned area in any particular year is expected to exceed this threshold with probability p, or in other words, this return level is expected to be exceeded on average once every $1/p$ years. For instance, p = 0.02 corresponds to a 50-year return period.

For GPD, the return level Z_p was determined by inverting eq. 2:

$$[5] \qquad z_p = \begin{cases} \dfrac{\sigma^*}{\xi}\left(p^{-\xi} - 1\right) \\ \sigma^* \ln\left(\dfrac{1}{p}\right) \end{cases}$$

It should be noted that the probability p here is adjusted to take into account the probability of an exceedance of the threshold (Coles 2001).

The Nelder–Mead method was applied to obtain the maximum likelihood estimation of each parameter in all three types of distributions, i.e., GEV, GPD, and PP. Advantages of the maximum likelihood estimation method have been illustrated in previous studies, including the flexibility it offers (Naveau et al. 2005), the easiness of incorporation of covariates (Coles and Dixon 1999), and its stability of parameter estimation (Brabson and Palutikof 2000). A quantile–quantile (q–q) plot was produced as a diagnostic check for the fitting of each distribution. In a q–q plot, less deviation from the 1:1 central tendency indicates a better fitting of the modeled outputs to observations. In a perfect case, all points lie in the 1:1 central tendency line, meaning that the model can exactly simulate the observations. Since the number of observations decreases when the burned area increases, the ability of the q–q plot to evaluate goodness-of-fit in the upper tails is reduced. However, the q–q plot is still a useful tool for checking if a model fitting is reasonable.

The free MATLAB package EVIM (Gençay et al. 2001) and package R version 2.12.1 and the “extRemes” package for R were used for statistical analyses (R Development Core Team 2009; Gilleland et al. 2009).

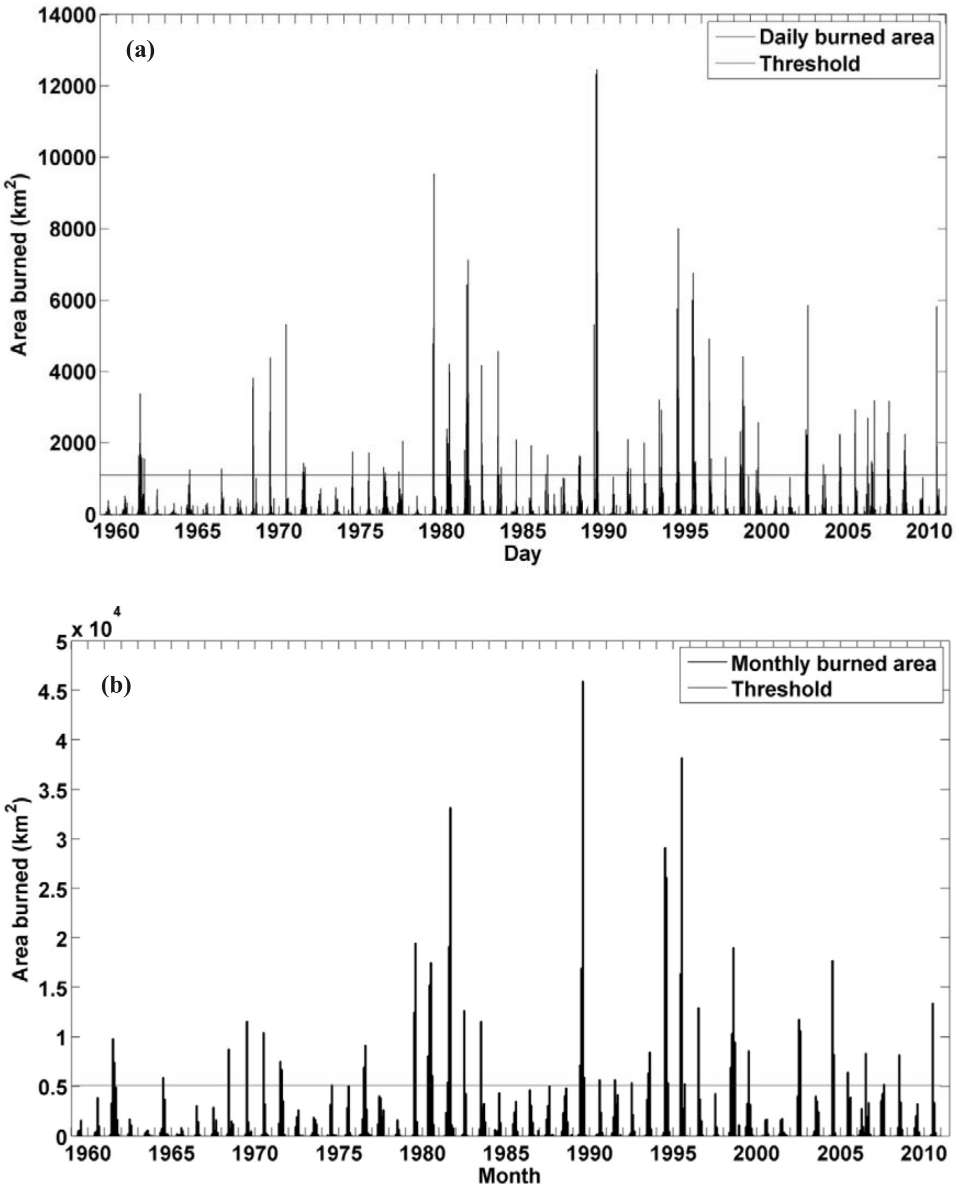
Modeling extreme fire events

For the whole and each portion of the data set, we firstly fitted the GEV distribution to the annual maxima of fire sizes of which the interannual variation is substantially large, e.g., in the whole Canadian data set (Fig. 2). Then, we fitted the GPD and PP distributions to the declustered data set with burned areas exceeding a predefined threshold u.

Using the extremal index to get the optimized run length and the 99% quantile as the original threshold, we declustered the whole data set and each data subset. For example, by assuming that exceedances belong to the same cluster if they are separated by less than 17 (run length) values below a given threshold u, we declustered the whole data set. After declustering, the extremal index is larger than 1, meaning that the filtered exceedances are independent. A data set with 65 effective exceedances was obtained, which accounts for approximately 21% of the total burned area in Canada during the period 1959–2010. Then, based on the sensitivity test of parameters and quantile changes among a range of thresholds, we determined the threshold for the GPD and PP models.

Similarly, we determined a threshold u = 130 km^2 to decluster the human-caused fires and we obtained 87 effective exceedances that account for 41% of the total human-caused burned areas. For lightning-caused fires, we defined 1050 km^2 as the threshold to obtain 64 filtered exceedances that account for 22% of the total lightning-caused burned areas. In our regional analysis, we defined a threshold u = 350 km^2 for the Boreal Shield and 187 km^2 for the Taiga

Fig. 2. Time series of Canadian (*a*) daily- and (*b*) monthly-summed burned areas from 1959 to 2010. In each panel, the horizontal line represents the threshold used in the peak over threshold method for the whole Canadian fire data set.



Shield and we obtained 75 and 64 effective exceedances, respectively. In the Boreal Shield, the effective exceedances resulted in 31% of the total burned area, while in the Taiga Shield, the exceedances accounted for 36% of the total burned area. In each declustered data set, all exceedances are statistically independent (i.e., extremal indexes are all larger than 1). Finally, we fitted the GPD and PP models to each of the declustered data sets.

A similar processing method was applied to the monthly data set but used a different definition of threshold. For the monthly data set, since there are only 624 data points, a 99% quantile threshold would lead to only six exceedances for the POT method. Thus, we used the 52 largest burned area fire data points as the exceedances (e.g., 5159 km^2 for the whole Canadian monthly data set shown in Fig. 2*b*). After declustering, we obtained 32, 37, 30, 30, and 35 effective exceedances for the Canadian fire, human-caused fire, lightning-caused fire, Boreal Shield fire, and Taiga Shield fire data set, respectively.

Results and discussion

The organized daily and monthly fire events show some commonalities in the extreme value statistics, i.e., fire size distributions are all heavy tailed (Tables 1 and 2). All three extreme distributions are well fitted to both daily and monthly fire events with reasonable confidence intervals. However, since the monthly extreme value analysis used much fewer data points, the standard errors or confidence intervals of estimates are relatively higher and wider than those in the daily extreme value analysis (Tables 1 and 2). Because most of fire prediction models (e.g., FWI system: Van Wagner 1987) are at a daily step, results of our extreme value analysis on daily-organized fire events could be easily compared with

Table 1. Maximum likelihood estimates of the parameters and corresponding fire return levels with standard errors (or confidence intervals) for the GEV, GPD, and PP models fitted to the time series of daily-organized fire records (km^2) from 1959 to 2010.

	Whole data set		Anthropogenic fires		Lightning fires		Boreal Shield		Taiga Shield	
Parameter	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)
GEV										
Location μ	1491.1	208.8	275.0	47.9	1313.4	202.3	766.0	133.7	270.0	55.6
Scale σ	1220.4	192.5	279.1	54.4	1133.4	194.0	752.9	122.2	326.4	62.7
Shape ξ	0.449	0.179	0.800	0.207	0.532	0.207	0.434	0.198	0.792	0.204
52-year return level	14713	(8295, 32242)	8086	(6516, 14171)	16529	(8410, 40266)	8638	(0, 19657)	9205	(0, 25444)
GPD										
Scale σ*	1254.1	284.4	261.0	55.6	1047.7	243.3	942.0	183.2	413.8	83.3
Shape ξ	0.321	0.194	0.570	0.194	0.435	0.203	0.144	0.157	0.467	0.168
52-year return level	12124	(4073, 20176)	5500	(347, 10652)	13323	(2984, 23662)	5999	(3022, 8975)	5466	(1359, 9573)
PP										
Location μ	1393.6	178.1	286.1	47.1	1277.4	151.6	704.4	130.3	277.2	59.6
Scale σ	1347.4	272.7	350.0	61.6	1146.2	244.6	993.6	156.8	455.8	88.6
Shape ξ	0.321	0.194	0.570	0.194	0.435	0.203	0.144	0.157	0.467	0.168
52-year return level	12078	(4086, 20070)	5477	(353, 10602)	13263	(3003, 23524)	5980	(3024, 8937)	5445	(1359, 9531)
Largest observation	12467		5325		12467		6347		11685	

Note: For each ecozone, we analyzed the combination of anthropogenic and natural fire data sets. The 95% confidence intervals for the 52-year return level were calculated using a profile likelihood method.

Published by NRC Research Press

Table 2. Maximum likelihood estimates of the parameters and corresponding fire return levels with standard errors (or confidence intervals) for the GEV, GPD, and PP models fitted to the time series of monthly-organized fire records (km^2) from 1959 to 2010.

	Whole data set		Anthropogenic fires		Lightning fires		Boreal Shield		Taiga Shield	
Parameter	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)	Estimate	1 SE (or 95% CI)
GEV										
Location μ	4582.9	615.0	564.4	89.5	3903.2	582.6	2119.9	375.6	554.1	114.0
Scale σ	3733.8	576.8	538.4	94.5	3546.8	565.9	2170.3	347.1	673.3	127.9
Shape ξ	0.470	0.151	0.668	0.180	0.530	0.162	0.363	0.191	0.782	0.200
52-year return level	47304	(27109, 99547)	10981	(0, 27272)	51219	(27454, 114451)	21153	(14583, 46757)	18469	(7640, 28777)
GPD										
Scale σ*	5943.6	1875.2	811.1	236.6	6915.2	2286.5	2530.5	619.6	741.4	252.7
Shape ξ	0.308	0.274	0.473	0.252	0.224	0.275	0.219	0.187	0.812	0.326
52-year return level	41443	(26451, 75098)	8273	(4681, 14807)	39461	(26163, 73234)	16347	(11321, 26469)	15889	(6723, 31018)
PP										
Location μ	2316.3	1441.5	380.5	143.6	934.3	2086.9	1519.2	433.7	527.4	135.8
Scale σ	5068.6.5	2339.0	681.9	248.2	6066.2	3168.6	2364.7	677.4	525.1	239.7
Shape ξ	0.309	0.287	0.472	0.252	0.224	0.303	0.218	0.187	0.812	0.326
52-year return level	41329	(12087, 70571)	8227	(1257, 15198)	39301	(14009, 64593)	16282	(7769, 24795)	15771	(4719, 36260)
Largest observation	45888		8450		45450		22425		22181	

Note: For each ecozone, we analyzed the combination of anthropogenic and natural fire data sets. The 95% confidence intervals for the 52-year return level were calculated using a profile likelihood method.

results from other models and potentially benefit forest management. In this study, we analyzed both the daily and monthly data sets, but we focused on presenting the results of daily extreme value analysis.

Three extreme statistical distributions were reasonably fitted to either the annual maximum fire sizes (GEV distribution) or fire sizes that exceeded certain thresholds (GPD and PP). All fittings are heavy tailed ($\xi > 0$), which agrees with Moritz's (1997) findings for a Californian fire analysis. Estimates of GEV distribution parameters (Table 1) indicate that the annual maxima of Canadian daily fire sizes are highly episodic (Fig. 2*a*). Similar findings were found for the four data subset analyses. Generally, both the magnitude and the variability of annual maximum lightning fires are larger than those of human-caused fires. Although the Boreal Shield ecozone experienced a higher magnitude of annual maximum fire size, its interannual variability is not as high as that of the Taiga Shield ecozone. However, because the annual maxima method only used 52 values in this case, the uncertainty of these estimates is considerably large.

Comparison among three extreme fire statistical methods

Since the GEV method used different data sources from those used in the GPD and PP methods, there is no need to compare their parameters directly, except the fire return levels. Among three methods, the GPD and PP methods are preferred. A main reason is that the GEV approach only takes into account one single value (annual maxima) per block (year), therefore neglecting a majority of available data (Katz et al. 2005; Blanchet et al. 2009). Directly fitting the GEV distribution to annual maxima over a short time period covered very few values, and in our case, only 52 values were accounted for. This could lead to a large uncertainty, especially when more than one extreme event exists in a block (year). In contrast, because the GPD and PP models use more available data, estimates are thus expected to be more accurate. Furthermore, forest managers are always interested in the occurrence of extremely large fires exceeding a certain threshold, not just the annual maximum fire sizes. Compared with the GEV approach, the return levels derived from the GPD and PP methods are more comparable with observations for either the whole Canadian fire data set or the four data subsets of ecozones or ignition sources. However, for using monthly data sets, GEV might be a priority method, since the number of exceedances of the monthly data points is only 52 in our study.

The premise to use a Poisson process is that the described events are independently and identically distributed. Since previous studies have shown that the medium and large fires are independently and identically distributed events (Malamud et al. 2005), a Poisson process could be used to represent the large fire events. Because the GPD and PP distributions are effectively equivalent (eq. 3), any inference made by the PP model could equivalently be made using the GPD method (Blanchet et al. 2009). It is generally considered better to model with the PP method because its parameter uncertainty is estimated simultaneously rather than orthogonally in the GPD (Coles 2001). However, due to its relative simplicity and the level of accuracy desired in this study, the GPD is preferred. In fact, the return levels derived from GPD and PP fittings are very close to each other (Tables 1 and 2). Thus, we mainly presented the results from the GPD fittings. We found that the GPD is reasonable in characterizing the exceedances in both the daily and monthly analyses (Fig. 3).

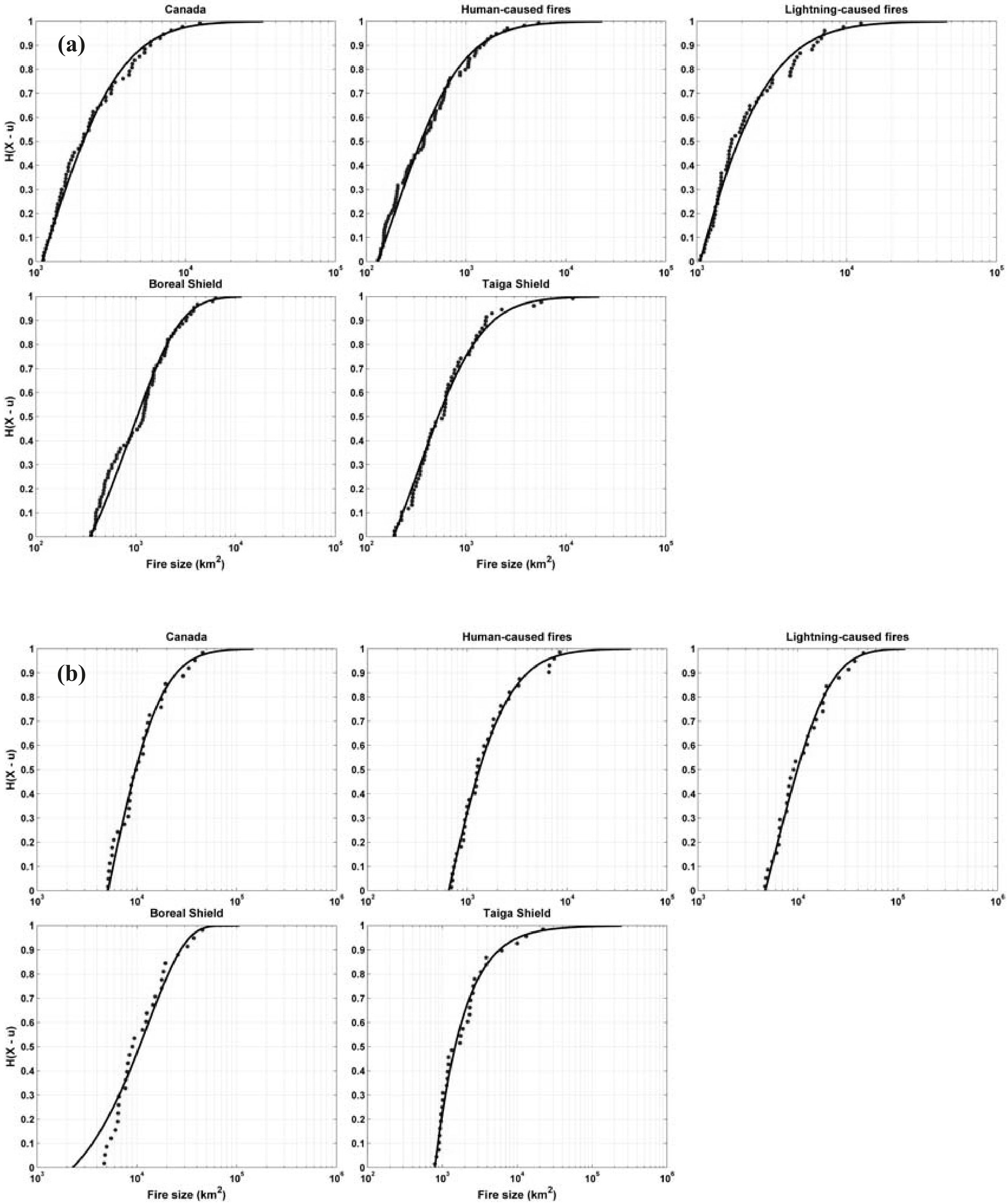
However, there are two major sources of uncertainty to our estimates using the GPD and PP models. The first source of uncertainty is related to the fixed threshold. In declustering, we applied the 99% quantile as the threshold to filter the data. The 99% quantile is a sufficiently high value that was determined based on the trade-off between bias and variance. A sensitivity test was then conducted to investigate how parameter estimates are sensitive to a range of thresholds (Fig. 4). Because the 99% quantile corresponds to fairly stable scale and shape parameters, we employed it as the threshold for the POT approach. The estimated return levels (e.g., Fig. 5 for the whole data set) are stable despite the subjective choice of threshold. Although some models are able to automatically account for the uncertainty due to threshold selection (e.g., Behrens et al. 2004; Tancredi et al. 2006), they still require some subjective assessment for the threshold choice. In this study, to quantify the uncertainty with a fixed threshold, we estimated the uncertainty range of the CDF for the fitted GPD model (Fig. 6). In future efforts, the threshold could be modeled as a time-varying variable to handle covariates (e.g., climates) for all GPD parameters (Tancredi et al. 2006).

The second source of uncertainty is the dependence between exceedances. The Canadian fire records (1959–2010) are best described as a nonstationary time series with some unknown dependence structure. However, there is no general theory established for this nonstationary process or providing likelihood to incorporate the cluster-induced dependence in the extremes. Although some previous studies provide methods to explicitly model the dependence structure using time series or covariates for model parameters, e.g., the GARCH model (Engle 2001), the model specification and threshold determination are still an unsolved problem (Pauli and Coles 2001; Bali and Weinbaum 2007; Zhao et al. 2011).

Nevertheless, as demonstrated in Coles (2001), the standard extreme value models (e.g., GPD) are still applicable in the presence of dependence. To deal with the dependence between threshold excesses, we declustered the fire data using the extremal index as the indicator of how much clustering of exceedances of a threshold occurs in the limit of the distribution. Before declustering, we find that all extremal indexes are less than 0.5, which means that there is some dependence in the limit. After declustering, all extremal indexes are equal to or larger than 1, which suggests that the declustered data points are independent. However, the deficiency of declustering is that we lost some useful information because the number of available data points decreases much after declustering. Consequently, the degrees of freedom decrease in calculating the confidence intervals.

Naturally, the extreme fire events could be related to other variables, referred to as a covariate (Coles 2001). For instance, the fire size could be substantially influenced by the wind speed. However, to explicitly model the dependence structure of extreme fire events is still an unsolved problem so far. Further efforts could model any combination of extreme fire model parameters as a function of time or other covariates (e.g., climate).

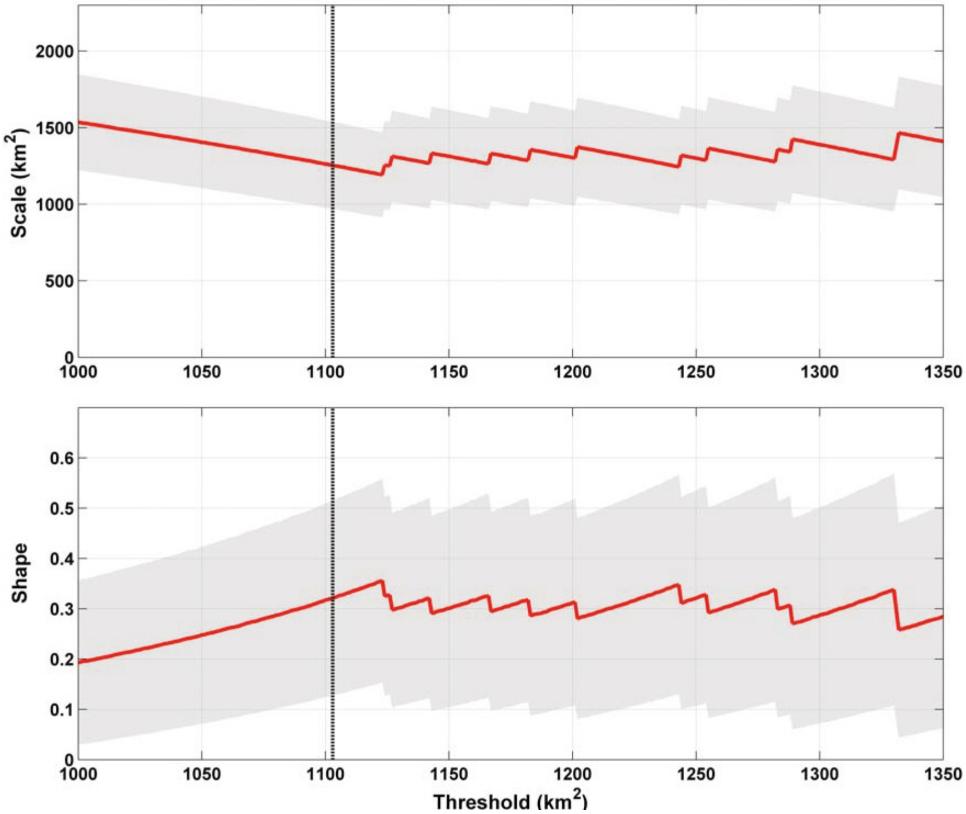
Fig. 3. Cumulative density functions of the estimated GPD models and the (*a*) daily- and (*b*) monthly-summed burned area exceedances over the predefined threshold *u*. The estimated models are plotted as curves and the actual data as circles. The *x*-axis is on a logarithmic scale.



Statistics of Canadian extreme fires

Based on the 65 effective exceedances, estimates for the two GPD parameters are 1254.1 km^2 (σ^*) and 0.321 (ξ) with a significant negative correlation (ρ = –0.71). This significant negative correlation exists in all analyses with different data sets. The scale parameter σ^*, which controls the spread

Fig. 4. Maximum likelihood estimates with uncertainty ranges of the modified scale and shape parameters in a GPD fitting to the exceedances in the whole Canadian fire data set over a range of thresholds. The uncertainty range is determined by ±1 SD. The vertical line shows the threshold (1103 km^2) that we used.



of the distribution, is not high given that all of the data points used are larger than 1103 km^2, while a GPD quantile is linear in σ^*. This implies that the characterization of extreme values is significantly governed by the shape parameter ξ in which the GPD quantile is nonlinear. Similar to other extreme fire studies (e.g., Moritz 1997 for California), the heavy tail is evident. A relatively small shape parameter (less than 0.5) implies that the extreme fire events have finite variance. In other words, there is definitely a finite upper bound on fire sizes. The maximum area of connected combustible forest patches or fire-stopping weather events (e.g., heavy rainfall) could be the limit. Consequently, this would contribute to a right-censored distribution of fire sizes.

The Gumbel case cannot be estimated because it is a single point in a continuous parameter space, so we tested against the null hypothesis that the shape parameter is actually zero (i.e., the exponential case, hence a light tail behavior). A likelihood ratio test (5% level) for $\xi = 0$ does not accept the exponential hypothesis and the p value is less than 0.001. In analysis of data subsets, similar likelihood ratio tests all reject the Gumbel case. An almost linear q–q plot (Fig. 7) indicates that the assumption for using the GPD is reasonable. In addition, Fig. 5 shows that estimated return levels are stable across a range of thresholds, which indicates that the fitted GPD model performs well in predicting the most extreme fire size within a specified return period, despite the uncertainty in the threshold selection.

Using the estimated parameters, we calculated the m-year return levels (Fig. 8) at the tail with a 95% confidence interval. We found that as the return periods increase, the increasing rate of return levels decreases, although the confidence interval becomes wider. One possible reason is that the environmental barriers (e.g., mountain, lake, wetlands) could be an important limitation for the spread of extreme fire events. For example, topographic barriers may slow the growth rates of large fires in the mountainous regions because fires would spread up into higher elevation areas, which are likely to be drier and with sparser vegetation, slowing the rate of growth and the fire would go out when a fire front reached a ridge (Reed and McKelvey 2002).

Statistics of anthropogenic and natural extreme fires

The fitted GPD for lightning fires has larger scale ($\sigma^* = 1047.7$ km^2) and lower shape ($\xi = 0.435$) parameters compared with those for human-caused fires ($\sigma^* = 261.0$ km^2, $\xi = 0.570$). Due to the significant correlation between σ^* and ξ, we investigated return levels instead to check the extremely large fire behaviors. The agreement in Fig. 9*a* between observations and our estimates suggests that the fitted GPD model works well.

We found that the estimated m-year return level of lightning-caused fires is always larger than that of human-caused fires, which suggests that lightning ignitions tend to cause more severe large fires. A two-sample t test supports that this difference is statistically significant. This difference could be the result, at least in part, of forest suppression. In particular, a high proportion of lightning fires tends to occur in remote areas and these regions are usually designated as

Fig. 5. Estimated return levels (quantiles) at a range of thresholds for the whole Canadian fire data set. The middle line shows the estimated value and the upper and lower lines represent the upper and lower bounds of the 95% confidence interval. The vertical lines shows the threshold (1103 km^2) that we used.

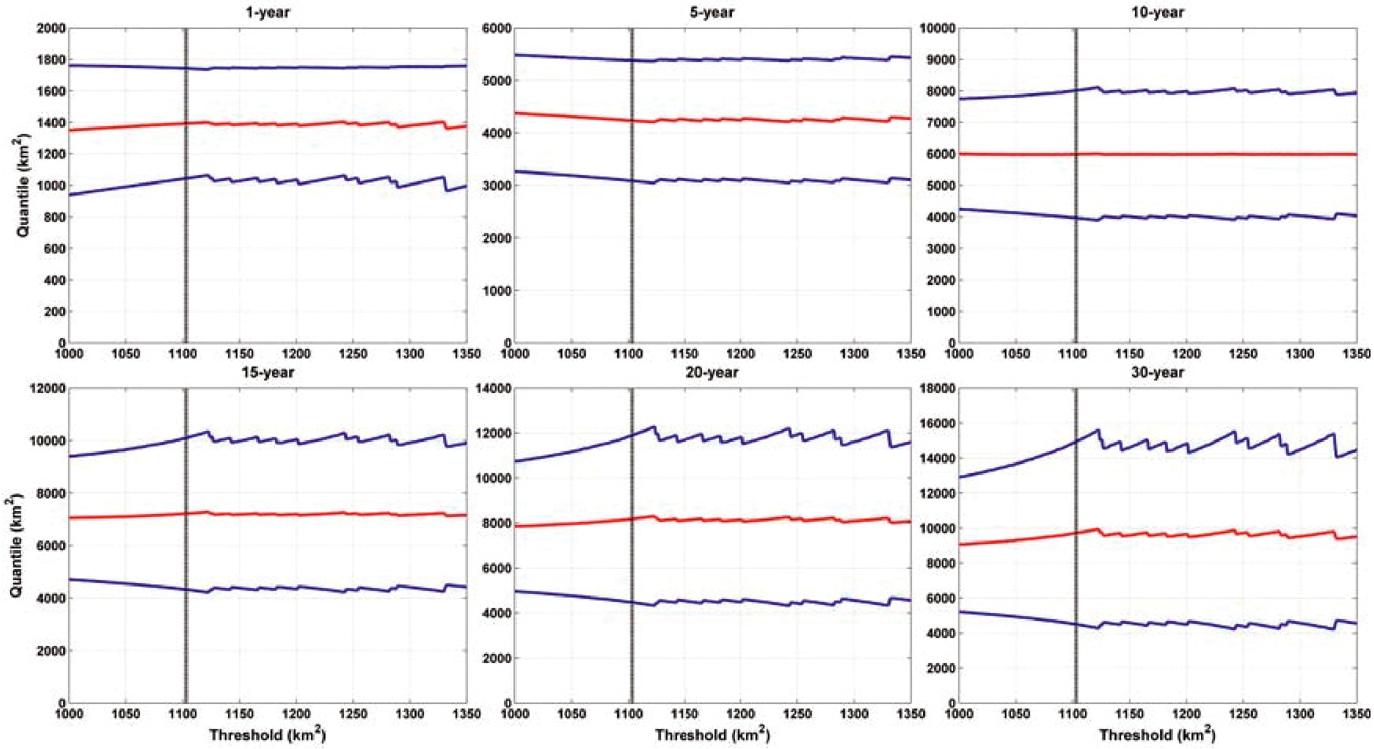
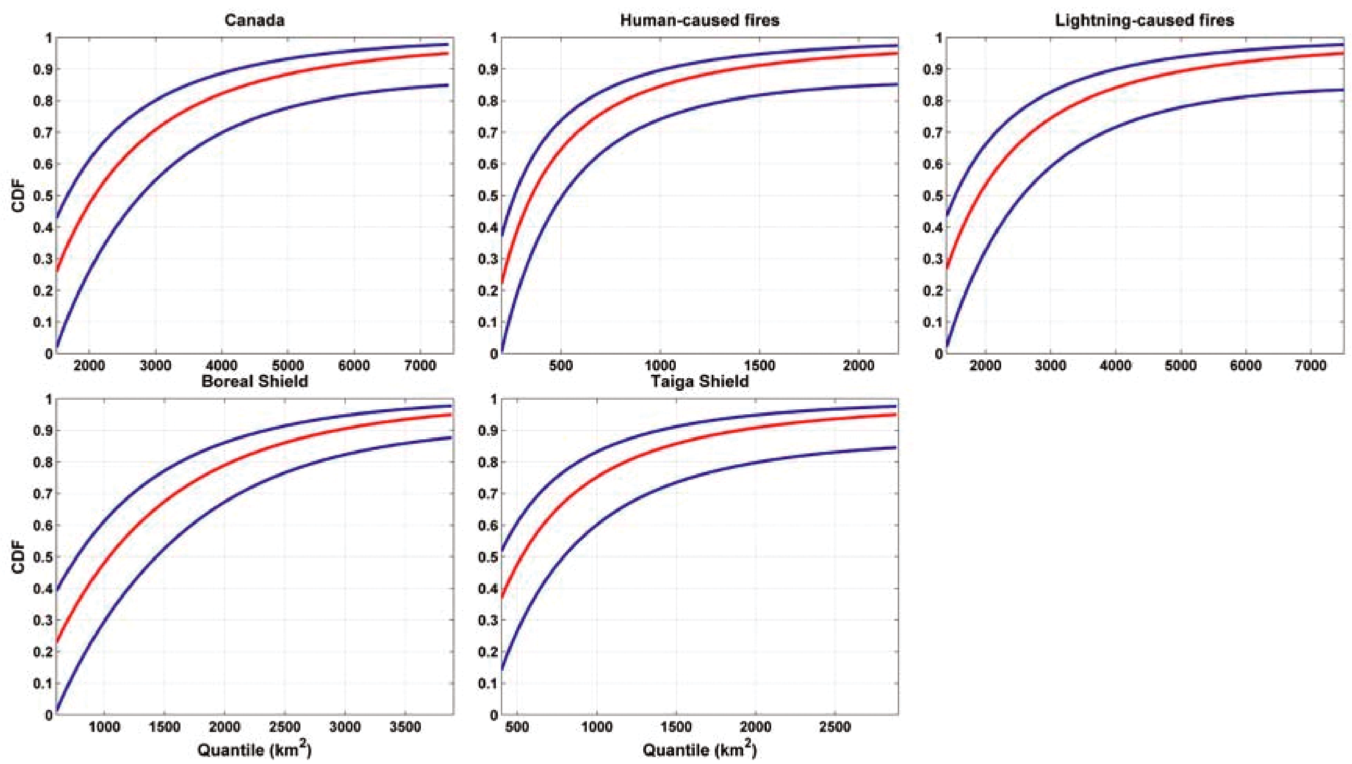
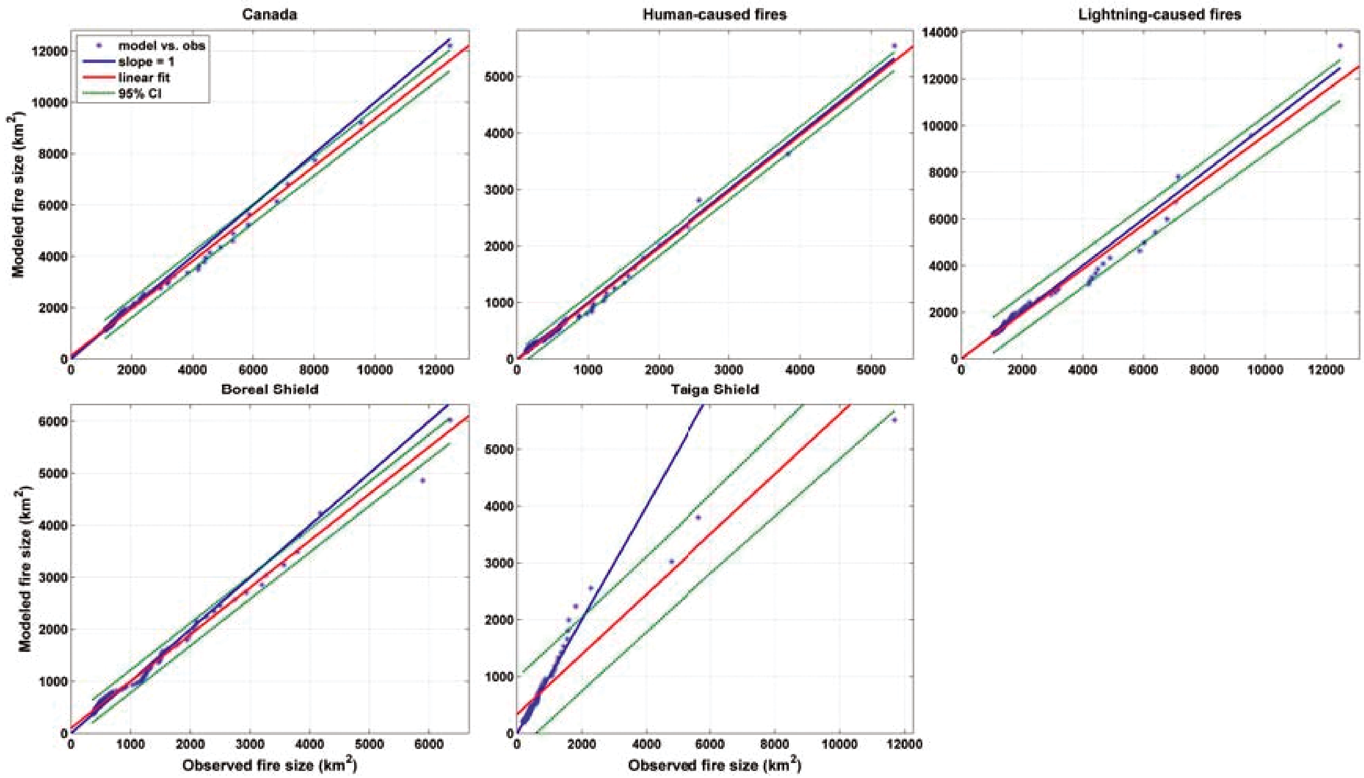


Fig. 6. Uncertainty range of the CDF for a range of quantiles with a fixed threshold in the GPD models. The upper and lower lines represent the upper and lower bounds of the 95% confidence interval.



"lower priority" zones that receive little or no fire suppression, since fires occurring there generally have little or no significant detrimental impact on public safety and forest values (Stocks et al. 2001). Very few fires in these regions are human-caused, as is shown in the map by Stocks et al. (2002) and Fig. 1. Because these remote regions do not have

Fig. 7. Quantile–quantile plots for the GPD fit to daily-organized burned area for (*a*) the whole of Canada, (*b*) human-caused and (*c*) lightning-caused fires, and fires in (*d*) the Boreal Shield ecozone and (*e*) the Taiga Shield ecozone. Systematic departures from the slope-1 line (dark line) suggest that the models are lacking good fit. The light line is the linear regression of the estimates against the observations with the 95% confidence interval for new observations (lines above and below).



aggressive fire suppression, the majority of fires are allowed to perform their natural functioning (Stocks et al. 2002) and therefore could reach an extremely large fire size. In addition, extreme fire weather (i.e., unusually high temperature with low rainfall) and fire-prone ecosystems (Stocks et al. 2002) in remote regions also make it easy for a fire to extend to an extremely large size. In contrast, human-caused fires occur in more populated areas and are usually reported more quickly and suppressed. Canada has effective fire suppression in populated areas and along the road network to limit the total burned area. For instance, Canada has an aggressive fire policy in the dense population regions near the Canada–US border where human-caused fires account for high percentages of fires. In these regions, anthropogenic fires can hardly achieve large sizes. Our findings are consistent with those of previous studies. Comparing fire sizes relative to levels of protection indicates that, on average, fires in the largely unprotected regions of the boreal zone are much larger than fires in intensively protected regions (Stocks 1991; Ward and Tithecott 1993; Stocks et al. 2001).

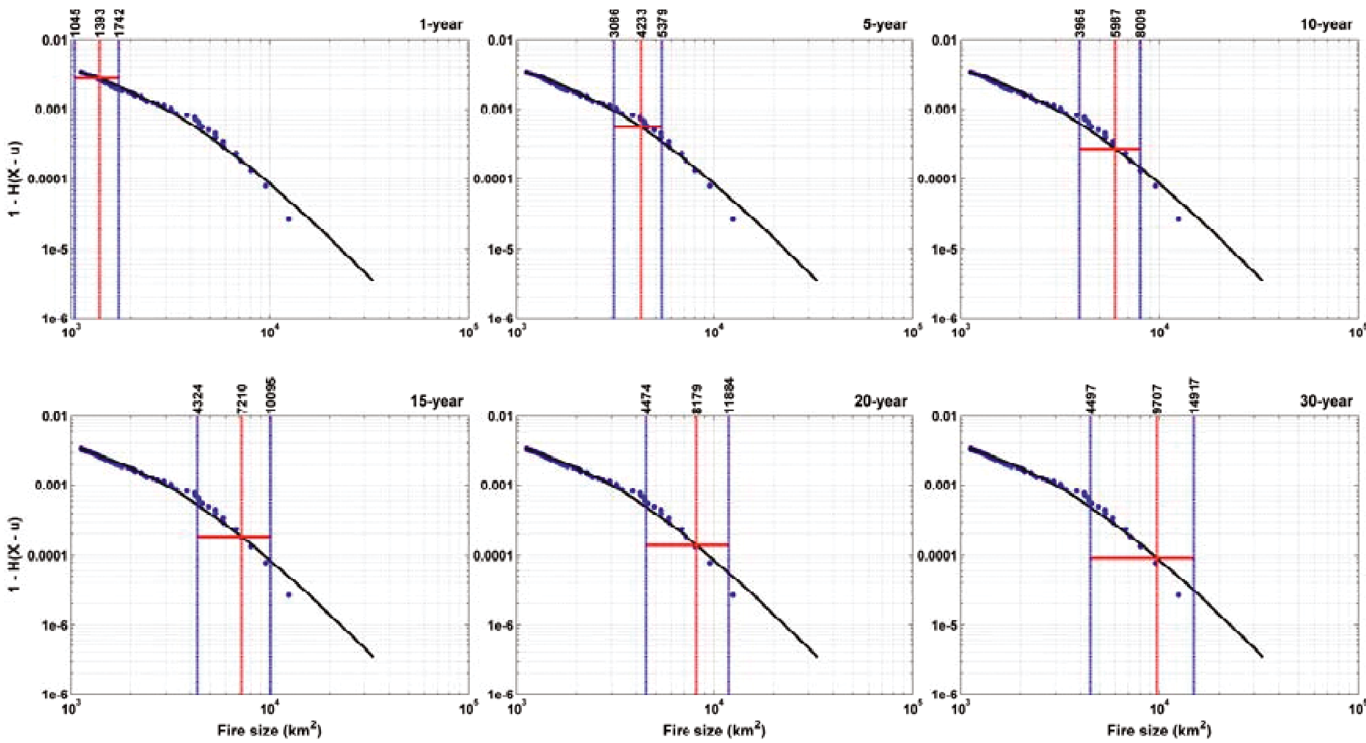
Based on the results presented above, we concluded that anthropogenic fires and natural fires exhibit substantially different statistics of extreme fire events, while they show some commonalities of extreme fire behaviors. The *q*–*q* plots for the fitted GPDs (Fig. 5) of both types of fire are approximately linear, meaning that the assumptions for using the GPD for both types of fire are reasonable. Although the departure increases as the burned area increases, the fitted values still show reasonable agreement with the observed values. Because lightning fires dominate the Canadian large fire behaviors (Stocks et al. 2002), the statistics of lightning-caused extreme fires are very similar to those of the whole data set.

Statistics for extreme fires in the Boreal Shield and Taiga Shield ecozones

The Boreal Shield and Taiga Shield ecozones both experienced a number of extreme fire events during the 52-year period. Due to different fire weather and fire regimes, the Boreal Shield ecozone could be subdivided into two divisions, west and east, separated by the Hudson Plains ecozone. The west subecozone has a strong continental climate and major fire activity, especially in the northern area. In the east subecozone, the fire climate is not as extreme as in the west division and fire suppression is aggressive. Due to similar reasons, the Taiga Shield ecozone could also be divided into east and west subecozones separated by the Hudson Bay. The two parts have similar features, while the east subecozone has a mild continental climate and it generally has less severe fire weather conditions (FAO 2002).

Estimates of parameters represent substantially different statistics of extreme fires between these two ecozones. As a diagnostic check, *q*–*q* plots for the fitted GPDs (Fig. 5) are almost linear, indicating that the assumed form of distribution is reasonable for the Boreal Shield ecozone. However, for the Taiga Shield ecozone, the *q*–*q* plots show that the GPD model performs poorly in estimating the most extreme fires. A higher scale parameter for Boreal Shield large fires

Fig. 8. Estimated *m*-year return levels and 95% confidence intervals for the Canadian daily-organized fire data based on the estimated GPD parameters. At the daily scale, the 1-, 5-, 10-, 15-, 20-, and 30-year return levels correspond to the 99.726, 99.945, 99.973, 99.982, 99.986, 99.991 percentile values. The estimated tail is plotted as a solid line and the actual data as circles. The vertical middle line is the estimated *m*-year return level with a lower bound (left line) and upper bound (right line).



(942.0 km^2) implies a higher spread than that for Taiga Shield fires (413.8 km^2). However, similar to the human-caused fires, a higher shape ξ of Taiga Shield fires (0.467 versus 0.144) could sufficiently increase the probabilities of large fires, especially the extremely large ones. Due to the dependence between parameters, we further investigated the return levels between fires in the two different ecozones. We found that the return level of Boreal Shield fires is always larger than that of Taiga Shield fires at the same return period. However, this difference decreases as the return period increases (Fig. 9*b*).

A smaller spread of extremely large fires is due to the widely covered wetlands in the Boreal Shield ecozone. In addition, the Boreal Shield east subecozone has extensive and successful fire suppression and the fire climate is not as extreme as in the west subecozone. In contrast, the Taiga Shield ecozone is dominated by coniferous forests and lichen woodlands; thus, the barriers due to vegetation coverage are not as critical as those in the Boreal Shield ecozone. The two subregions of the Taiga Shield ecozone are dominated by lightning fires and most fires are allowed to burn naturally. Thus, it is easier for fires to spread to an extremely large size.

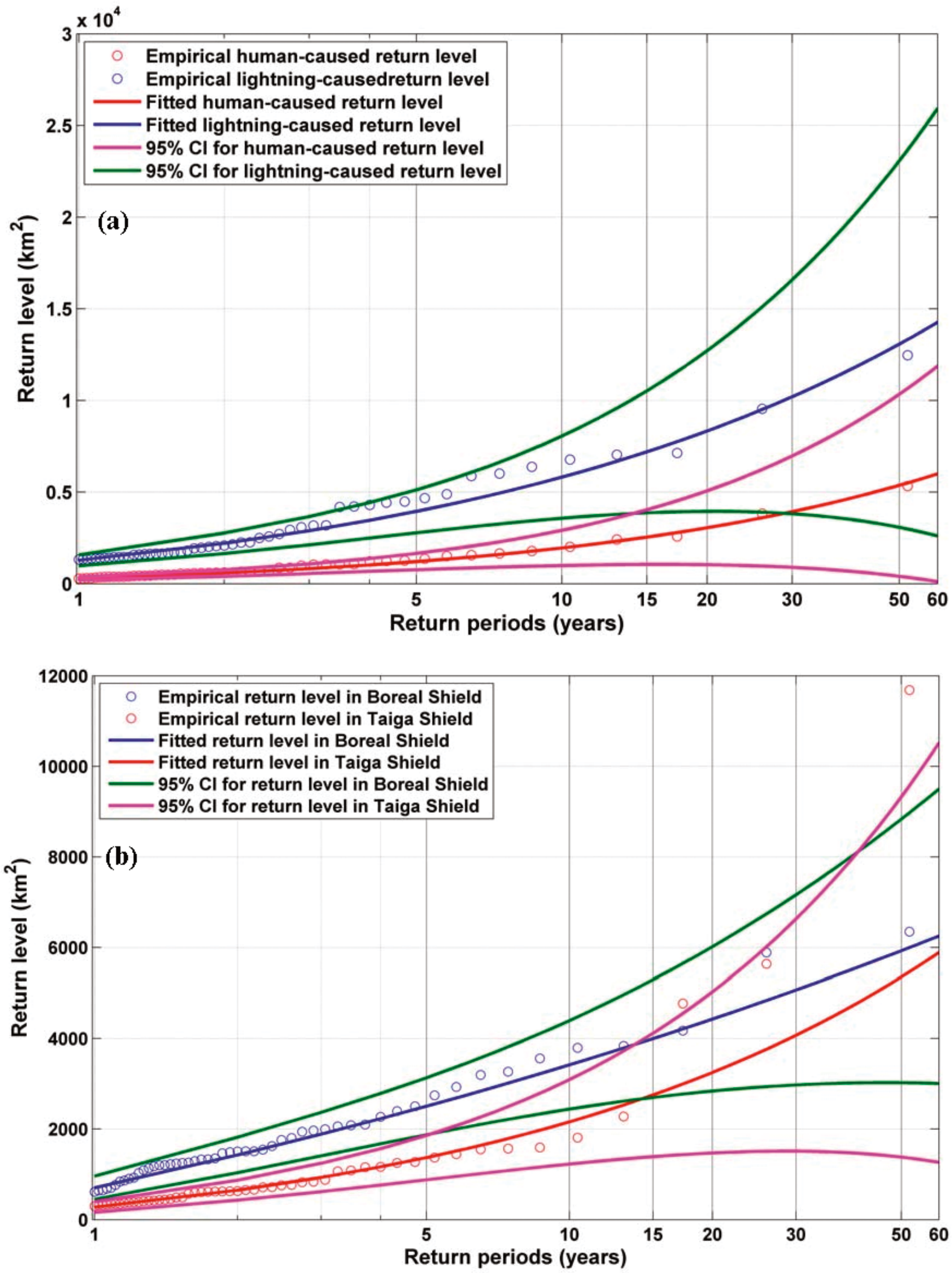
For the Boreal Shield ecozone, our estimated *m*-year return levels all show good agreement with historical observations. However, for the Taiga Shield ecozone, when the return period is longer than 15 years, the corresponding return levels show a large difference from the observations (Fig. 9*b*). For both ecozones, the large confidence intervals obtained for extreme return levels imply that there is little information with which to make future predictions with any degree of certainty. In different ecozones, there is a distinct limit on the fire size, which functions as an upper cutoff for return levels of wildfires. For example, topography (i.e., mountain or lake) or vegetation (i.e., wetlands) barriers could constrain the most extreme fire size.

Implications of EVT to wildfire study and management

Large fires have played a major role in Canadian boreal forest dynamics. Although Canadian provincial and territorial agencies have much improved their fire management systems, forest fires continue to affect the Canadian forest resource (Stocks et al. 2001). Extreme fires always lead to substantially social, economic, and ecological consequences and suppression is costly (Beverly and Martell 2005). The large fires across the southern interior of the province of British Columbia in 2003, including the Okanagan Mountain Park forest fire that burned parts of the city of Kelowna, is an example (Woolford et al. 2010). Since extreme fires are responsible for the majority of burned areas, characterization of these events is important to better understand fire regimes.

Traditional studies focusing on mean and variance of fire sizes poorly characterize extreme fire events due to their rare occurrences. The power–law frequency–area models (e.g., Cumming 2001; Malamud et al. 2005; Jiang et al. 2009) are successful in describing relationships between fire frequency and burned area for medium and large fires. However, the shape of the tail at extreme values in fire size distribution is still unclear. The extreme statistical approaches have proven to perform better in characterizing the upper tail occupied by extreme values, thereby providing better estimations of ex-

Fig. 9. Empirical return levels and estimated return levels with 95% confidence intervals for (*a*) human-caused and lightning-caused fires and (*b*) Boreal Shield and Taiga Shield fires.



treme fire return levels, which could be used to evaluate the potential fire risk in terms of frequency of a specific size. EVT could be a useful tool for forest managers when assessing fire hazard risk because it provides a flexible and justifiable approach for extrapolating the tail distribution.

Previous studies (e.g., Fosberg et al. 1996; Stocks et al. 1998) also presented an increasing trend of extreme fire danger across Canada under a warming climate. Apart from the many and diverse issues surrounding the fidelity and recording methods of Canadian fires (1959–2010), we demonstrated that stationarity is the least likely state of the extreme fire events. However, there is no developed extreme theory for this real nonstationary series (Coles 2001). Nevertheless, the extremal models presented in this study are still able to be used as a basic template of which parameters could be further modeled as functions of time series or some covariates (e.g., climate), although how climate change impacts fire regimes is still unknown (Woolford et al. 2010). Characterizing extreme fire events and quantifying the fire risks associated with climate changes are needed with the EVT method. Quantification of extreme fire return levels provides a way to estimate the risk of extreme fires for a region. This knowledge and information should be helpful for timber supply modeling and analyzing the vulnerability of forest-based communities (Beverly and Martell 2005). The extreme statistics also provide useful information for stakeholders concerning the size of the largest fires and their expected return levels.

Conclusions

Large fires are a major disturbance in Canadian forests and affect both the climate system and ecosystems. To date, it is still a challenge to predict the large fire size and frequency. Here, we used three extreme statistical distributions to characterize large fires. All fitted distributions are heavy tailed (Fréchet or Pareto distribution). Among the three distributions, GPD and PP both perform more reasonably in describing the upper tail of the large fire distributions than the GEV method does. In Canada, the Boreal Shield ecozone has a higher risk of extreme large fire occurrence than the Taiga Shield. Human-caused and lightning-caused fires exhibit very different extreme behaviors and lightning fires dominate the extreme fire regime in the whole of Canada. Future work in the identification of daily-organized fire events is greatly needed for the daily time series analysis of extreme fire events. The statistical properties of extreme fires described here should be useful for assessing fire risks associated with regional factors or ignition sources for fire and landscape managers. The information of fire return levels derived from these extreme value distributions should also be valuable for studying fire dynamics and their impacts on the climate and ecosystems in this region.

Acknowledgements

The wildfire database is courtesy of Canadian provincial, territorial, and federal fire management agencies. We would like to thank two reviewers for their constructive comments on an earlier version of the manuscript. This study is supported by the NSF Arctic System Science Program (NSF-0554811), NSF Carbon and Water in the Earth Program (NSF-0630319), NASA Land Use and Land Cover Change Program with project NASA- NNX09AI26G and Department of Energy funded project DE-FG02-08ER64599.

References

ACIA. 2004. Impacts of a warming arctic: arctic climate impact assessment. Cambridge University Press, New York. Available at http://www.acia.uaf.edu.

Alvarado, E., Sandberg, D.V., and Pickford, S.G. 1998. Modeling large forest fires as extreme events. Northwest Sci. **72**(Special Issue): 66–75.

Bali, T.G., and Weinbaum, D. 2007. A conditional extreme value volatility estimator based on high-frequency returns. J. Econ. Dyn. Control, **31**(2): 361–397. doi:10.1016/j.jedc.2005.10.002.

Behrens, C.N., Lopes, H.F., and Gamerman, D. 2004. Bayesian analysis of extreme events with threshold estimation. Stat. Model. **4**(3): 227–244. doi:10.1191/1471082X04st075oa.

Beverly, J.L., and Martell, D.L. 2005. Characterizing extreme fire and weather events in the Boreal Shield ecozone of Ontario. Agric. For. Meteorol. **133**(1–4): 5–16. doi:10.1016/j.agrformet.2005.07.015.

Blanchet, J., Marty, C., and Lehning, M. 2009. Extreme value statistics of snowfall in the Swiss Alpine region. Water Resour. Res. **45**(5): W05424. doi:10.1029/2009WR007916.

Brabson, B.B., and Palutikof, J.P. 2000. Tests of the generalized Pareto distribution for predicting extreme wind speeds. J. Appl. Meteorol. **39**(9): 1627–1640. doi:10.1175/1520-0450(2000)039<1627:TOTGPD>2.0.CO;2.

Burton, P.J., Parisien, M., Hicke, J.A., Hall, R.J., and Freeburn, J.T. 2008. Large fires as agents of ecological diversity in the North American boreal forest. Int. J. Wildland Fire, **17**(6): 754–767. doi:10.1071/WF07149.

Canadian Forest Service. 2010. National fire database — agencyfire data. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alta. Available at http://cwfis.cfs.nrcan.gc.ca/en_CA/nfdb/poly.

Coles, S.G. 2001. An introduction to statistical modeling of extreme values. Springer, New York.

Coles, S.G., and Dixon, M.J. 1999. Likelihood-based inference for extreme value models. Extremes, **2**(1): 5–23. doi:10.1023/A:1009905222644.

Cumming, S.G. 2001. A parametric model of the fire-size distribution. Can. J. For. Res. **31**(8): 1297–1303. doi:10.1139/x01-032.

Davison, A.C., and Smith, R.L. 1990. Models for exceedances over high threshold. J. R. Stat. Soc. B, **52**(3): 393–442.

Dey, D.C., and Guyette, R.P. 2000. Anthropogenic fire history and red oak forests in south-central Ontario. For. Chron. **76**(2): 339–347.

Engle, R.F. 2001. GARCH 101: the use of ARCH/GARCH models in applied econometrics. J. Econ. Perspect. **15**(4): 157–168. doi:10.1257/jep.15.4.157.

FAO. 2002. Guidelines on fire management in temperate and boreal forests. Forest Protection Working Papers, Working Paper FP/1/E. Forest Resources Development Service, Forest Resources Division. FAO, Rome.

Ferro, C.A.T., and Segers, J. 2003. Inference for clusters of extreme values. J. R. Stat. Soc. Ser. A Stat. Soc. **65**: 545–556.

Flannigan, M.D., Bergeron, Y., Engelmark, O., and Wotton, B.M. 1998. Future wildfire in circumboreal forests in relation to global warming. J. Veg. Sci. **9**(4): 469–476. doi:10.2307/3237261.

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., and Stocks, B.J. 2005. Future area burned in Canada. Clim. Change, **72**(1-2): 1–16. doi:10.1007/s10584-005-5935-y.

Fosberg, M.A., Stocks, B.J., and Lynham, T.J. 1996. Risk analysis in strategic planning: fire and climate change in the boreal forest. *In* Fire in ecosystems of boreal Eurasia. *Edited by* J.G. Goldammer and V.V. Furyaev. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 495–505.

Gençay, R., Selçuk, F., and Ulugülyagci, A. 2001. EVIM: a software package for extreme value analysis in MATLAB. Stud. Nonlinear Dyn. Econometr. **5**(3): 213–239. doi:10.1162/10811820160080103.

Gilleland, E., Katz, R., and Young, G. 2009. extRemes: extreme value toolkit. R package, version 1.60. Available at http://cran.r-project.org/web/packages/extRemes/index.html.

Gillett, N.P., Weaver, A.J., Zwiers, F.W., and Flannigan, M.D. 2004. Detecting the effect of climate change on Canadian forest fires. Geophys. Res. Lett. **31**(18): L18211. doi:10.1029/2004GL020876.

Guyette, R.P., Muzika, R.M., and Dey, D.C. 2002. Dynamics of an anthropogenic fire regime. Ecosystems (N.Y.), **5**: 472–486.

Jiang, Y., Zhuang, Q., Flannigan, M.D., and Little, J.M. 2009. Characterization of wildfire regimes in Canadian boreal terrestrial ecosystems. Int. J. Wildland Fire, **18**(8): 992–1002. doi:10.1071/WF08096.

Kasischke, E.S., and Turetsky, M.R. 2006. Recent changes in the fire regime across the North American boreal region — spatial and temporal patterns of burning across Canada and Alaska. Geophys. Res. Lett. **33**(9): L09703. doi:10.1029/2006GL025677.

Katz, R.W., Brush, G.S., and Parlange, M.B. 2005. Statistics of extremes: modeling ecological disturbances. Ecology, **86**(5): 1124–1134. doi:10.1890/04-0606.

Koenker, R. 2005. Quantile regression. Cambridge University Press, Cambridge, U.K.

Krawchuk, M.A., and Cumming, S.G. 2009. Disturbance history affects lightning fire initiation in the mixedwood boreal forest: observations and simulations. For. Ecol. Manage. **257**(7): 1613–1622. doi:10.1016/j.foreco.2009.01.019.

Krawchuk, M.A., Cumming, S.G., Flannigan, M.D., and Wein, R.W. 2006. Biotic and abiotic regulation of lightning fire initiation in the mixedwood boreal forest. Ecology, **87**(2): 458–468. doi:10.1890/05-1021. PMID:16637370.

Kushla, J.D., and Ripple, W.J. 1997. The role of terrain in a fire mosaic of a temperate coniferous forest. For. Ecol. Manage. **95**(2): 97–107. doi:10.1016/S0378-1127(97)82929-5.

Leadbetter, M.R., Lindgren, G., and Rootzén, H. 1983. Extremes and related properties of random sequences and processes. Springer-Verlag, New York.

Malamud, B.D., Millington, J.D.A., and Perry, G.L.W. 2005. Characterizing wildfire regimes in the United States. Proc. Natl. Acad. Sci. U.S.A. **102**(13): 4694–4699. PMID:15781868. doi:10.1073/pnas.0500880102.

Moritz, M.A. 1997. Analyzing extreme disturbance events: fire in Los Padres National Forest. Ecol. Appl. **7**(4): 1252–1262. doi:10.1890/1051-0761(1997)007[1252:AEDEFI]2.0.CO;2.

Nash, C.H., and Johnson, E.A. 1996. Synoptic climatology of lightning-caused forest fires in subalpine and boreal forests. Can. J. For. Res. **26**(10): 1859–1874. doi:10.1139/x26-211.

Naveau, P., Nogaj, M., Ammann, C., Yiou, P., Cooley, D., and Jomelli, V. 2005. Statistical methods for the analysis of climate extremes. C. R. Geosci. **337**(10–11): 1013–1022. doi:10.1016/j.crte.2005.04.015.

Pauli, F., and Coles, S.G. 2001. Penalized likelihood inference in extreme value analyses. J. Appl. Stat. **28**(5): 547–560. doi:10.1080/02664760120047889.

Podur, J., Martell, D.L., and Knight, K. 2002. Statistical quality control analysis of forest fire activity in Canada. Can. J. For. Res. **32**(2): 195–205. doi:10.1139/x01-183.

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at http://www.r-project.org.

Reed, W.J., and McKelvey, K.S. 2002. Power-law behaviour and parametric models for the size-distribution of forest fires. Ecol. Model. **150**(3): 239–254. doi:10.1016/S0304-3800(01)00483-5.

Rigby, R.A., and Stasinopoulos, D.M. 2005. Generalized additive models for location, scale and shape. Appl. Stat. **54**: 507–554.

Schoenberg, F.P., Peng, R., Huang, Z., and Rundel, P. 2003. Detection of non-linearities in the dependence of burn area on fuel age and climate variables. Int. J. Wildland Fire, **12**(1): 1–6. doi:10.1071/WF02053.

Smith, R.L. 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Stat. Sci. **4**(4): 367–377. doi:10.1214/ss/1177012400.

Stocks, B.J. 1991. The extent and impact of forest fires in northern circumpolar countries. *In* Global biomass burning: atmospheric, climatic, and biospheric implications. *Edited by* J.S. Levine. MIT Press, Cambridge, Mass. pp. 197–202.

Stocks, B.J., Fosberg, M.A., Lynham, T.J., Mearns, L., Wotton, B.M., Yang, Q., Jin, J.Z., Lawrence, K., Hartley, G.R., Mason, J.A., and McKenny, D.W. 1998. Climate change and forest fire potential in Russian and Canadian boreal forests. Clim. Change, **38**(1): 1–13. doi:10.1023/A:1005306001055.

Stocks, B.J., Wotton, B.M., Flannigan, M.D., Fosberg, M.A., Cahoon, D.R., and Goldammer, J.G. 2001. Boreal forest fire regimes and climate change. *In* Remote sensing and climate modeling: synergies and limitations. *Edited by* M. Beniston and M.M. Verstraete. Springer, New York. pp. 233–246.

Stocks, B.J., Mason, J.A., Todd, J.B., Bosch, E.M., Wotton, B.M., Amiro, B.D., Flannigan, M.D., Hirsch, K.G., Logan, K.A., Martell, D.L., and Skinner, W.R. 2002. Large forest fires in Canada, 1959–1997. J. Geophys. Res. **107**(8149). doi:10.1029/2001JD000484.

Tancredi, A., Anderson, C., and O'Hagan, A. 2006. Accounting for threshold uncertainty in extreme value estimation. Extremes, **9**(2): 87–106. doi:10.1007/s10687-006-0009-8.

Van Wagner, C.E. 1987. Development and structure of the Canadian Forest Fire Weather Index. For. Tech. Rep. 35. Canadian Forestry Service, Petawawa National Forestry Institute, Ottawa, Ont.

Ward, P.C., and Tithecott, A.G. 1993. Impact of fire management on the boreal landscape of Ontario. Publ. 305. Ministry of Natural Resources, Aviation, Flood and Fire Management Branch, Sault Ste. Marie, Ont.

Woolford, D.G., Cao, J., Dean, C.B., and Martell, D.L. 2010. Characterizing temporal changes in forest fire ignitions: looking for climate change signals in a region of the Canadian boreal forest. Environmetrics, **21**(7-8): 789–800. doi:10.1002/env.1067.

Wotton, B.M., Martell, D.L., and Logan, K.A. 2003. Climate change and people-caused forest fire occurrence in Ontario. Clim. Change, **60**(3): 275–295. doi:10.1023/A:1026075919710.

Xiao, J., and Zhuang, Q. 2007. Drought effects on large fire activity in Canadian and Alaskan forests. Environ. Res. Lett. **2**(4): 044003. doi:10.1088/1748-9326/2/4/044003.

Zhao, X., Scarrott, C., Oxley, L., and Reale, M. 2011. GARCH dependence in extreme value model with Bayesian inference. Math. Comput. Simul. **81**(7): 1430–1440. doi:10.1016/j.matcom.2010.08.002.

Zhuang, Q., Melillo, J.M., Sarofim, M.C., Kicklighter, D.W., McGuire, A.D., Felzer, B.S., Sokolov, A., Prinn, R.G., Steudler, P.A., and Hu, S. 2006. CO_2 and CH_4 exchanges between land ecosystems and the atmosphere in northern high latitudes over the 21st century. Geophys. Res. Lett. **33**(17): L17403. doi:10.1029/2006GL026972.