Available online at www.sciencedirect.com

ScienceDirect

Agriculture and Agricultural Science Procedia 1 (2010) 251–257

Agriculture and Agricultural Science Procedia

International Conference on Agricultural Risk and Food Security 2010

Modeling Agricultural Catastrophic Risk

XU Lei[a, b]， ZHANG Qiao[a, b] *

[a]Agricultural Information Institute，Chinese Academy of Agricultural Sciences，Beijing 100081
[b]Key Laboratory for Agricultural Early Warning based on Intelligent Technology and Systems，Ministry of Agriculture, Beijing 100081

Abstract

The paper aims to develop approaches for modeling agricultural catastrophic risk. According to extreme value theory, this study applies the block maxima and peak-over-threshold models to analyze and evaluate the agricultural catastrophic risk illustrated by cases of extreme rainfall in Jilin Province. This study provides the basic statistical analysis framework for evaluating agricultural catastrophic risk.

Key words：agricultural catastrophic risk；extreme value theory；BMM model；POT model

* Corresponding author. Tel.: +86-010-82109651
E-mail address: xulei2005@caas.net.cn

2210-7843 © 2010 Published by Elsevier B.V.
doi:10.1016/j.aaspro.2010.09.031

1. Introduction

Agricultural catastrophic risks are related to extreme events that have low-probability (cannot be easily predicted), but have relatively serious negative agricultural economic consequences. Assessing agricultural catastrophic risk involves determining the probability of extreme events normally seen in weather disasters and the expected level at which these events occur. More importantly, interest generally lies in the measure of extreme event PDF underlying events that trigger loss. Thus, the basic concept and the key step of modeling agricultural catastrophic risk for the purpose of assessing it is fully analogous to modeling the probability distribution for the weather events in question. Traditional statistics mostly focuses on laws governing average. However, catastrophic events can lead to severe consequences; thus, these events fall under specific distributions. Modeling extreme risk using traditional statistics approaches, therefore, is potentially misleading or biased. A previous study shows that the extreme value model has an advantage in modeling and assessing extreme risk compared with traditional parametric methods.
The paper is structured as follows. Section 2 provides a quick overview of extreme value theory (EVT) and approaches for modeling agricultural catastrophic risk. Section 3 discusses the application of the block maxima model (BMM) and peak-over-threshold (POT) model for the analysis and evaluation of the agricultural catastrophic risk illustrated by examples of extreme rainfall in Jilin Province. Section 4 provides a general overview of the results and reveals potential directions for further research.

2. Extreme Value Theory and Modeling

EVT (Fisher and Tippett, 1928) dates back to the late 1920s and has been extensively applied in many subjects during the last several decades. It can potentially provide a promising solution to modeling extreme risk because it is primarily concerned with quantifying the stochastic behavior of a process at its largest or smallest, or the events over a threshold. Generally, there are two principal approaches in modeling extreme values: the BMM and POT models.

2.1. Block Maxima Model

The BMM focuses on the statistical behavior of the largest or smallest value in a sequence of independent random variables. Assume breaking up a sequence into blocks of size n (with n reasonably large), and extracting only the maximum observation $M_i(i=1,2...n)$ from each block. M_n is normalized to obtain a non-degenerated limiting distribution, known as the generalized extreme value distribution (GEV), with c.d.f: $G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$,

defined on the set $\{z:1+\xi(z-\mu)/\sigma>0\}$, where μ, σ > 0, and ξ are location, scale, and shape parameters, respectively. Note that Gumbel distribution, Frechet distribution, and Weibull distribution correspond to the cases ξ=0, ξ > 0, and ξ < 0, respectively. Then, the log-likelihood for the GEV parameters is given by the equation $l(\xi,\sigma,\mu) = -n\log\sigma - (1+1/\xi)\sum_{i=1}^{n}\log\left[1+\xi(\frac{M_i-\mu}{\sigma})\right] - \sum_{i=1}^{n}\left[1+\xi(\frac{M_i-\mu}{\sigma})\right]^{-1/\xi}$. The maximization of this equation with respect to the parameter vector (ξ,σ,μ) leads to the maximum likelihood estimate (MLE) with respect to the entire GEV family (Gumbel 1958; Castillo 1988).

2.2. Peak-over-Threshold Model

The POT approach is based on the generalized Pareto distribution in the following manner (Pickands, 1975). POT methods use a more natural way of determining whether an observation is extreme. All values greater than the given high value (threshold) are considered. Previous studies have shown that if the block maxima has an approximate distribution of GEV, then threshold excesses have a corresponding generalized Pareto distribution (GPD). Thus, the excess values above a high level as with c.d.f: $H(y) = 1-(1+\frac{\xi y}{\sigma})^{-1/\xi}$, defined on the set $(Y = X - u)$, where σ (σ >0) and ξ ($-\infty < \xi < +\infty$) are scale and shape parameters, respectively. Then, the log-likelihood for the GPD parameters is given by the equation $l(\sigma,\xi) = n\log\xi - n\log\sigma - (1+1/\xi)\sum_{i=1}^{n}\log(1+\xi y_i/\sigma)$. Thus, maximum likelihood procedures (MLE) can also be utilized to estimate the GPD parameters, given the threshold.

3. Empirical analysis and case study

3.1 The case of BMM model

This analysis is based on the series of annual maximum rainfall recorded for the typical crop areas in Jilin Province, China, from June to September (the most important months for crop production) for the period 1956–2007. Figure 1 shows that it seems reasonable to assume that the pattern of variation has stayed constant over the observation period. Thus, we model the data as independent observations from the GEV distribution.

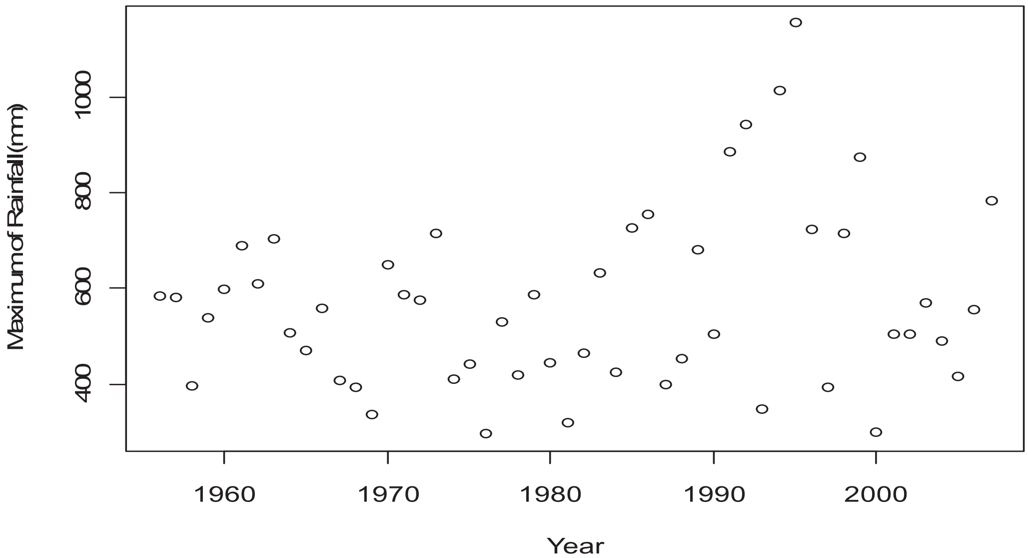

FIGURE 1. Scatter plot of annual maximum rainfall of Jilin Province

Maximization of the GEV log-likelihood for these data leads to the following estimate: $(\hat{\xi},\hat{\sigma},\hat{\mu}) = (0.03501458, 139.41185422, 482.8848686)$. We determine the type of the limiting distribution for maximum rainfall in Jilin Province, which is a type of Frechet distribution with the following form:

$$G(z) = \exp\left\{-\left[1+0.03501458\left(\frac{z-482.8848686}{139.41185422}\right)\right]^{-1/0.03501458}\right\}.$$

Figure 2 shows the various diagnostic plots for assessing the accuracy of the GEV model fitted to the maximum rainfall data. Neither the probability plot nor the quantile plot presents cause to doubt the validity of the fitted model: each set of plotted points is near-linear. The return level curve converges asymptotically to a finite level as a consequence of the positive estimate, although the estimate is close to zero and the respective estimated curve is close to a straight line. The curve also provides a satisfactory representation of the empirical estimates, especially once sampling variability is taken into account. Finally, the corresponding density estimate seems consistent with the histogram of the data. Consequently, all four diagnostic plots provide support to the fitted GEV model.

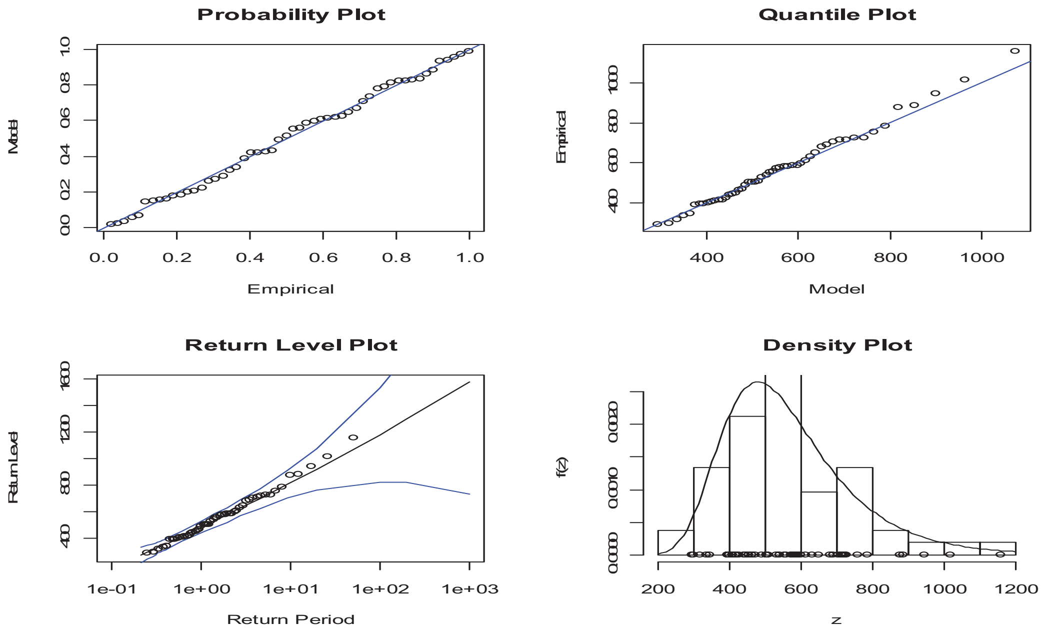

FIGURE 2 Diagnostic Plots for GEV fit to the Annual maximum rainfall of Jilin province

Assessing catastrophic risk is fully analogous to modeling the probability distribution for the extreme rainfall in question. In Table 1, we provide the exceedance probability intervals for the standard extreme value distribution. From the table, we can see that the highest probability of catastrophic rainfall is 25.14%, which has an interval between 400 mm and 500 mm.

Table 1 Risk evaluation for the extreme rainfall based on BMM model

P(300mm<rainfall<400mm)	P(400mm<rainfall<500mm)	P(500mm<rainfall<600mm)
13.97%	25.14%	23.32%
P(600mm<rainfall<700mm)	P(700mm<rainfall<800mm)	P(800mm<rainfall<900mm)
15.69%	9.11%	4.95%
P(900mm<rainfall<1000mm)	P(1000mm<rainfall<1100mm)	P(1100mm<rainfall)
2.64%	1.39%	1.62%

3.2 The case of POT model

We use the same rainfall data, but now consider the total daily data, which is above 233 mm. In the POT model, the determination of the threshold u is crucial. A threshold that is too low is likely to violate the asymptotic basis of the model and lead to bias; a threshold that is too high will generate minimal observations to estimate the parameters of the tail distribution function. This will also cause a high variance. We make use of the fact that if the GPD is the correct model for all the exceedances x_i above a given high threshold u_0, then the mean excess, i.e., the mean value of $(x_i - u)$, plotted against $u > u_0$, should give a linear plot (Davison and Smith, 1990) [Because $E[X_i - u_0]$ is a linear function of $u : u > u_0$]. By producing such a plot for values of u starting at zero, we can select reasonable candidate values for u_0. Mean residual of life plot for rainfall turns out reasonably well for all the excesses above $u_0 = 440$.

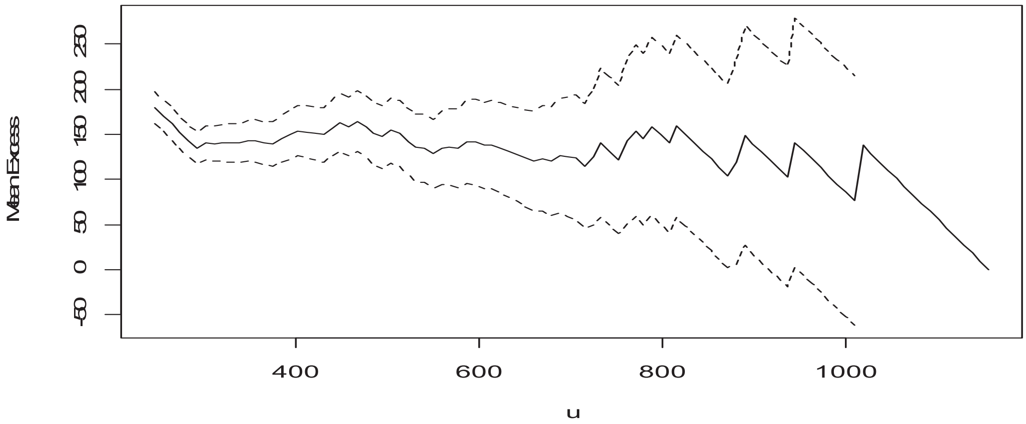

FIGURE 3 Mean residual of life plot for rainfall

We investigate the stability of our estimates of ξ and σ^*. Vertical lines in Figure 4 show 95% confidence intervals, which helped us assess the correct choice for the threshold $u_0 = 440$.

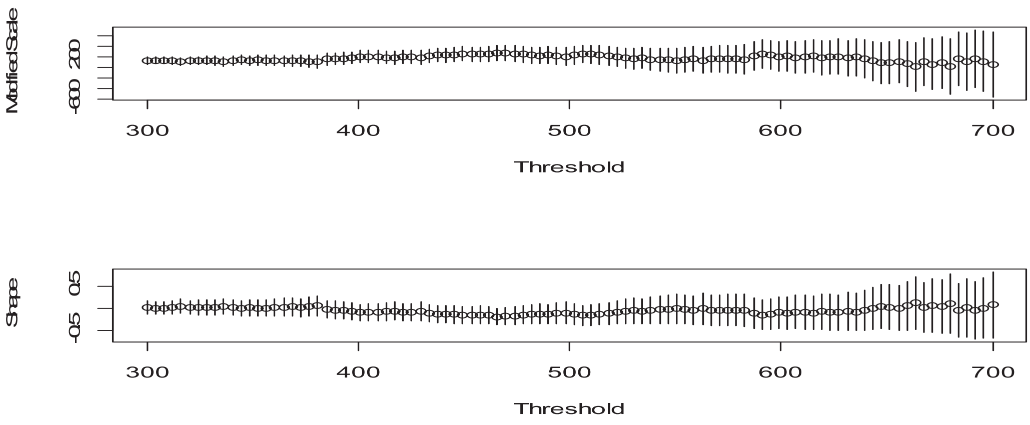

FIGURE 4 Parameter stability plots for rainfall

Once the threshold u is determined, we can estimate the parameters of GPD using MLE approach. The parameters and distribution function of GPD are $\sigma_u = 175.0396984$; $\xi = -0.1187855$. Thus,

$$H(x) = 1 - (1 - 0.1187855 \frac{x}{175.0396984})^{\frac{1}{0.1187855}}.$$

Figure 5 shows the various diagnostic plots for assessing the accuracy of the GPD model fitted to the rainfall data. Neither the probability plot nor the quantile plot presents cause to doubt the validity of the fitted model.

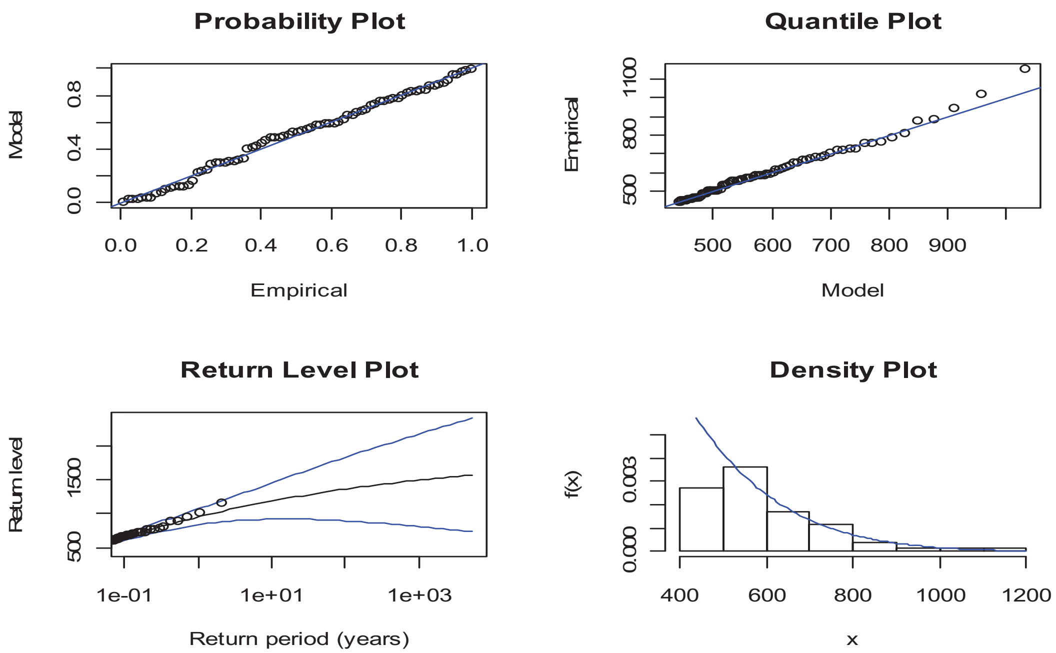

FIGURE 5 Diagnostic Plots for GPD fit to the daily rainfall of Jilin province

In the same way, we evaluate risk for extreme rainfall according to the probability distribution for extreme rainfall. From Table 2, we can see that the highest probability of catastrophic rainfall is 2.00%, which has an interval between 440 mm and 500 mm.

Table 2 Risk evaluation for the extreme rainfall based on POT model

P(440mm<rainfall<500mm)	P(500mm<rainfall<600mm)	P(600mm<rainfall<700mm)
2.00%	1.83%	0.79%
P(700mm<rainfall<800mm)	P(800mm<rainfall<900mm)	P(900mm<rainfall<1000mm)
0.30%	0.10%	0.03%
P(1000mm<rainfall<1100mm)	P(1100mm<rainfall)	
0.01%	9.61E-06	

4. Discussion and Outlook

The BMM and POT models are two effective approaches for evaluating agricultural catastrophic risks caused by extreme weather events. Because BMM only considers the largest or smallest events (the most common implementation of this approach for catastrophic weather data is to take block size to be one year), it is an inefficient approach if other data on the tail are available and of interest, and too narrow to be applied to a wide range of problems. Applications of these methods to cases of extreme rainfall in Jilin Province have shown that the predicted risk values obtained by the POT method are, in general, significantly below the corresponding predictions obtained using the traditional BMM. The POT approach can compensate for such shortcomings and be used to model all large (small) observations that exceed (fall below) a high (low) threshold.

Further research is directed to the more complex problem of agricultural catastrophic risk. To some extent, agricultural catastrophic risk is the consequence of extreme weather events. However, agricultural catastrophic risk is not the same as extreme weather risk. Other factors such as environment, agricultural investment, and farmer management should be taken into account. This means that the distribution for potential damages and losses after a certain type of extreme weather condition should be considered. We need to investigate the relationship between extreme weather events and catastrophic loss, and then we can evaluate the agricultural catastrophic risk using EVT and other relevant approaches.

Acknowledgement

This work was supported by the Key Projects of National Key Technology R&D Program—"Support System and Demonstration of Key Technologies of Intelligent-Analysis and Early-Warning for Security of Agricultural Products" (2009BADA9B00).

References

1. Fisher R. A. and Tippett L. H. C., Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Philos. Soc., 24 (1928)180–190.
2. Gumbel, E., Statistics of Extremes. Columbia University Press: New York, 1958.
3. Castillo, E., Extreme Value Theory in Engineering. Academic Press: San Diego, 1988.
4. Pickands, J., Statistical inference using extreme order statistics. The Annals of Statistics 3 (1975) 119–131.
5. Davison, A. C. and Smith, R. L., Models for exceedances over high thresholds (with discussion). J. R. Statist. Soc. B, 52 (1990) 393–442.