Return Periods and Return Levels Under Climate Change

Daniel Cooley
Department of Statistics
Colorado State University, Fort Collins, CO USA

cooleyd@stat.colostate.edu
(970) 491-5269

December 22, 2011

Abstract

We investigate the notions of return period and return level for a nonstationary climate. We discuss two general methods for communicating risk. The first eschews the term return period and instead communicates yearly risk in terms of a probability of exceedance. The second extends the notion of return period to the non-stationary setting. We examine two different definitions of return period under non-stationarity. The first, which appears in Olsen et al. (1998), defines the m-year return level as the level for which the expected waiting time until the exceedance is m years. The second, which appears in Parey et al. (2007) and Parey et al. (2010), defines the m-year return level as the level for which the expected number of events in an m year period is one. We illustrate the various risk communications with an application to annual peak flow measurements for the Red River of the North.

Keywords: risk communication, non-stationarity, floods

1 Introduction

1.1 Return Periods and Return Levels Under Stationarity

In many disciplines, return levels and return periods are used to describe and quantify risk. Classical work in probabilistic hydrology risk, precipitation frequency analysis, and other fields assumes a stationary climate. There is a general consensus in the scientific community that climate change has accelerated over the past few decades and that climate will continue to change in the coming decades primarily due to anthropogenic modifications of the Earth's atmosphere (IPCC, 2007). Consequently, there is growing interest to consider and account for non-stationarity when assessing risk. However, the assumption of stationarity has pervaded all areas of the above disciplines, including even the basic terminology. Two fundamental terms to these disciplines are return level and return period. These terms are relatively easily understood given stationarity, but become ambiguous in the non-stationary setting. The aim of this chapter is to explore these concepts under non-stationarity.

Because we will illustrate by analyzing an annual maximum series in Section 3, let us define the m-year return level as the high quantile for which the probability that the annual maximum exceeds this quantile is $1/m$. Under an assumption of stationarity, the return level is the same for all years, and this gives rise to the notion of the return period. Under stationarity, the return period of a particular event is the inverse of the probability that the event will be exceeded in any given year. Thus, the m-year return level is associated with a return period of m years. A return period can be *interpreted* in different ways, and we will explore different interpretations shortly. For now, let us simply acknowledge that in the stationary case there is a one-to-one relationship between a return level (the quantile) and a return period (the associated time interval).

The above definition in which return period is defined in terms of the *annual maximum's* probability of exceedance is not universal. Mays (2001, p. 317) equates the definition of return period with "average recurrence interval"; that is, the time between exceedance events. The difference between the two definitions arises because of the probability of having more than one exceedance in a given year. Because of the ambiguity of the definition of return period, NOAA's latest precipitation atlas effort (Bonnin et al., 2004) avoids using the term return period altogether. NOAA, like Mays, uses the terms average recurrence interval (ARI) and "annual exceedance probability" (AEP) to speak of return period as we defined it above. NOAA notes that the difference between the return levels associated with ARI and AEP is noticeable only for time intervals shorter 20 years, as the probability of multiple exceedances of very high thresholds in any year becomes negligible (Bonnin et al., 2004, Section 3.2). For the purposes of discussing the implications of non-stationarity in this chapter, the definitions of return period and return level in terms of the annual maximum are sufficient.

Due to the one-to-one relationship between return level and return period, in the stationary case it is straightforward to solve for either quantity given the value of the other. Design criteria may specify that a structure be built to withstand the (say) 100-year event, or regulations may require the designation of the 100-year floodplain, and the corresponding return level can be found for this period. A less common calculation is to find the return period associated with a particular level. An individual could wish to calculate the return period associated with his or her particular structure being flooded, or after a large event has occurred, one might wish to ascertain the event's associated return period.

1.2 Statistical Models for the Distribution's Tail

In practice, one does not know the exact distribution of the individual events or annual maximum, and one must estimate the upper tail of a distribution to assess risk. From the estimated distribution one can calculate the return period or return level of interest and additionally provide measures of uncertainty.

In the application in Section 3, we will rely on probabilistic results from extreme value theory to model the distribution of the annual maximum observation. There is a rich literature which describes the fundamental probability results of extreme value theory as well as the resulting statistical practice for modeling the tail derived from these results. Some recent references are de Haan and Ferreira (2006), Beirlant et al. (2004), and Coles (2001). Let $\{X_t\}$ denote a time series of our quantity of interest, for example, daily maximum streamflow or daily total precipitation. Let $M_n = \max_{t=1,\ldots,n} X_t$. The foundational result of extreme value theory states that if X_t are iid, and M_n can be linearly renormalized in such a way that its distribution converges as n grows, then it will converge to an extreme-value (equivalently, max-stable) distribution (Fisher and Tippett, 1928; Gnedenko, 1943). Further theoretical results state that the distribution of exceedances $P(X_t > x + u | X_t > u)$ should be well approximated by a generalized Pareto (GP) distribution above a sufficiently high threshold u (Balkema and De Haan, 1974; Pickands, 1975). If $\{X_t\}$ are not independent but are still identically distributed, the asymptotic distributions do not change, so long as certain relatively weak mixing conditions are met (Leadbetter et al., 1983).

In practice, an extreme value statistical analysis extracts a subset of data deemed extreme, and then fits a theoretically-justified distribution to this data subset. One approach is to construct a time series of block (e.g. annual) maxima, and another approach is to construct a subset of data which exceed a previously defined threshold (deemed a partial-duration series in hydrology). When the block-size n is fixed and large enough to assume that the asymptotic results provide a good approximation, the above theory suggests modeling M_n with a generalized extreme value (GEV) distribution. Threshold exceedance data can be modeled with GP distribution or an equivalent point process representation (Davison and Smith, 1990). Traditional extreme value results assume the data are identically distributed, but methods have been developed enable construction of non-stationary models based on the GEV and GP distributions or the point process representation of threshold exceedances (Katz, 2011; Smith, 1989).

The GEV and GP distributions' advantage is their asymptotic justification which is particularly useful for extrapolating the tail beyond the range of the data. However, risk analyses are not always based on distributions derived from extreme value theory. The log-Pearson type III distribution is widely used in hydrology and its use is sometimes mandated by government agencies (Vogel and Wilson, 1996). This chapter's discussion of return levels and periods ris not inherently tied to the GEV distribution that is used in Section 3; any model for the distribution's tail can be used.

1.3 Interpretations of Return Periods Under Stationarity

In the stationary case there is a one-to-one relationship between the m-year return level and m-year return period which is defined implicitly as the reciprocal of the probability of an exceedance in any one year. Return periods were assumedly created for the purpose of interpretation: a 100-year event may be more interpretable by the general public than a 0.01 probability of occurrence in any particular year. But this implicit definition gives rise to at least two interpretations of "an m-year event". The first interpretation is that the expected waiting time until the next exceedance is m years. The second is that the expected number of events in m years is 1. We show below that both of these interpretations are correct under a stationary assumption.

Let M_y denote the random variable representing the annual maximum for year y. Note that we have omitted the notational dependence on block size n and now include a year index y. For now, assume $\{M_y\}$ are iid with distribution function F. Given a return period of interest m, we can solve the equation

$$F(r_m) = P(M_y \leq r_m) = 1 - 1/m$$

for r_m, the associated return level. The first interpretation of the m-year event is the expected waiting time until an exceedance occurs. Let T be the year of the first exceedance. One recognizes

$$\begin{aligned} P(T = t) &= P(M_1 \leq r_m, M_2 \leq r_m, \ldots, M_{t-1} \leq r_m, M_t > r_m) \\ &= P(M_1 \leq r_m)P(M_2 \leq r_m) \ldots P(M_{t-1} \leq r_m)P(M_t > r_m) \\ &= P(M_1 \leq r_m)^{t-1} P(M_1 > r_m) \\ &= F^{t-1}(r_m)(1 - F(r_m)) \\ &= (1 - 1/m)^{t-1}(1/m), \end{aligned}$$

where the second line follows from the independence assumption and the third from stationarity. T is a geometric random variable, and it is well known that its expected value is m. That is, the expected waiting time for an m-year event is m years.

An alternative interpretation of an m-year event is that the expected number of events in m years is 1. Let N be the expected number of exceedances in m years; that is $N = \sum_{y=1}^{m} I(M_y > r_m)$ where I is the indicator function. Each year can be viewed as a trial, and since we have assumed M_y are iid, N has a binomial distribution:

$$P(N = k) = \binom{m}{k} (1/m)^k (1 - 1/m)^{m-k}. \tag{1}$$

It is straightforward to show that the expected value of N is 1.

1.4 Outline

The outline for the remainder of this chapter is as follows. In the next section, we discuss several options for communicating risk in the non-stationary case. In Section 3, we fit a stationary model and some non-stationary models to a time series of annual maximum river flows and convey the different risk measures for these various models. We conclude with a discussion in Section 4 where we touch on both implications of non-stationary models and expressions of risk.

2 Communicating Risk Under Non-stationarity

In this section we outline two general ideas for conveying risk when the process is assumed to be non-stationary. We start by explicitly conveying the changing nature of the risk by expressing the yearly probability of exceedance given a fixed level, or conversely the yearly exceedance level for a fixed probability of exceedance. Later, we extend the notion of return period to the non-stationary setting via both the expected waiting time and the expected number of events interpretations of return period. Throughout Section 2, we keep the independence assumption, but drop the assumption that the sequence $\{M_y\}$ is identically distributed.

2.1 Communicating Changing Risk

The most straightforward approach to communicating risk in the non-stationary setting is to give yearly estimates of risk. However, the idea of a yearly return period seems illogical, and it makes more sense to communicate risk via probabilities. Let F_y denote the distribution function of M_y. In any particular year, there still exists a one-to-one correspondence between a probability of exceedance and a high quantile.

Given a particular level of interest r, it is straightforward to express yearly risk in terms of probability. Letting $p(y) = P(M_y > r) = 1 - F_y(r)$, once F_y is estimated it is simple to provide yearly point estimates of the probability of an exceedance $p(y)$. These point estimates could be supplemented by employing the delta method (Casella and Berger, 2002, Section 5.5.4) to obtain a confidence interval.

Generally risk calculations proceed in the opposite direction: one starts with a measure of risk (e.g., a return period in the stationary case) and finds the corresponding level. Simply inverting the procedure in the previous paragraph allows one to start with a probability of exceedance p and solve $F_y(r_p(y)) = 1 - p$. As the exceedance level $r_p(y)$ changes with year, it clearly conveys to the user the changing nature of risk. Again, there are existing methods to obtain confidence intervals to convey uncertainty associated with $r_p(y)$

2.2 Return Periods and Return Levels Under Non-stationarity

Structures are designed for certain return periods based on the economic and human impacts of their failure. The dynamic risk valuations in Section 2.1 must be extended to make lifespan calculations in the non-stationary setting. Below we extend the expected waiting time and expected number of events interpretations of return period given in Section 1.3 to the non-stationary setting.

2.2.1 Return Period as Expected Waiting Time

One interpretation of return period given in Section 1.3 is the expected waiting time until an exceedance occurs. Olsen et al. (1998) use this definition to compute an m-year return level for non-stationary time series and compare this to the $1/m$ probability-of-exceedance level described above. Olsen et al. (1998) illustrate the difference in the definitions using simulated data. Here, we extend the basic definition set forth in Olsen et al. (1998) and discuss computational aspects. Specifically, we aim to find the level r_m for which the expected waiting time for an exceedance of this level is m years.

Let T be the waiting time (from $y = 0$) until an exceedance over a general level r occurs. Starting as before,

$$\begin{aligned}
P(T = t) &= P(M_1 \le r)P(M_2 \le r) \dots P(M_{t-1} \le r)P(M_t > r) \\
&= \prod_{y=1}^{t-1} F_y(r)(1 - F_t(r)) \\
\Rightarrow E[T] &= \sum_{t=1}^{\infty} t \prod_{y=1}^{t-1} F_y(r)(1 - F_t(r)) \qquad (2) \\
&= 1 + \sum_{i=1}^{\infty} \prod_{y=1}^{i} F_y(r),
\end{aligned}$$

where (2) is the definition that appears in Olsen et al. (1998) and the last line results from expanding the sum in the previous line and then collecting terms (shown in the appendix).

Defining the m-year return level r_m as the level which the expected waiting time until an exceedance occurs is m years, then r_m is the solution to the equation

$$m = 1 + \sum_{i=1}^{\infty} \prod_{y=1}^{i} F_y(r_m). \tag{3}$$

Because (3) cannot be written as a geometric series, solving for r_m is not straightforward. However, in the case that $F_y(r)$ is monotonically decreasing as $y \to \infty$ (that is, the extremes are getting more extreme), it is possible to bound the right-hand side of (3). For any positive integer L,

$$\begin{aligned} m &= 1 + \sum_{i=1}^{L} \prod_{y=1}^{i} F_y(r_m) + \sum_{i=L+1}^{\infty} \prod_{y=1}^{i} F_y(r_m) \\ \Rightarrow m &> 1 + \sum_{i=1}^{L} \prod_{y=1}^{i} F_y(r_m). \end{aligned} \tag{4}$$

Furthermore,

$$\begin{aligned} m &= 1 + \sum_{i=1}^{L} \prod_{y=1}^{i} F_y(r_m) + \prod_{y=1}^{L} F_y(r_m) \sum_{i=L+1}^{\infty} \prod_{y=L+1}^{i} F_y(r_m) \\ m &\leq 1 + \sum_{i=1}^{L} \prod_{y=1}^{i} F_y(r_m) + \prod_{y=1}^{L} F_y(r_m) \sum_{i=L+1}^{\infty} (F_{L+1}(r_m))^{i-L} \\ &= 1 + \sum_{i=1}^{L} \prod_{y=1}^{i} F_y(r_m) + \prod_{y=1}^{L} F_y(r_m) \frac{F_{L+1}(r_m)}{1 - F_{L+1}(r_m)}, \end{aligned} \tag{5}$$

where the inequality follows from the fact that $F_{L+1} \geq F_y$ if $y > L+1$. One can achieve bounds to any desired width by choosing L to be large enough. Solving for r_m must be done numerically, but this is done relatively easily since m is monotonically increasing with r_m.

2.2.2 Return Period as Expected Number of Events

The other interpretation of an m-year return period given in Section 1.3 is that the expected number of events in m years is one. Parey et al. (2007) and Parey et al. (2010) extend this definition to the non-stationary case.

We aim to find the level r_m for which the expected number of exceedances in m years is one. Let N be the number of exceedances that occur in the m years beginning with year $y = 1$ and ending with year $y = m$. As the probability of an exceedance is no longer constant from year to year, the distribution of N is no longer binomial. Letting

I be an indicator variable and r be a general threshold level, one obtains

$$\begin{aligned} N &= \sum_{y=1}^{m} I\left(M_y > r\right) \\ \Rightarrow E[N] &= \sum_{y=1}^{m} E[I\left(M_y > r\right)] \\ &= \sum_{y=1}^{m} P(M_y > r) \\ &= \sum_{y=1}^{m} (1 - F_y(r)). \end{aligned} \tag{6}$$

Setting (6) equal to one and solving, we define the m-year return level r_m to be the solution to the equation

$$1 = \sum_{y=1}^{m} (1 - F_y(r_m)). \tag{7}$$

3 Illustrative Example: Red River at Halstad

We examine annual peak flow measurements from the Red River of the North at Halstad, Minnesota, USA (station #05054500). Data were obtained from the US Geological Survey's National Water Information System.[1] The Halstad site has annual maximum data from 1942-2010. Additionally, we include the highest reading for 2011 (61600 cfs, 7:00 am, April 12) from the site's real-time data archive[2]. The Red River was selected because it has experienced recent newsworthy floods in 1997, 2009, and 2011. In Section 4, we further discuss possible implications of the selection of the Red River.

Figure 1 shows the annual maximum flow measurements at Halstad. There is some indication that the annual maximum flows have increased over this period. The highest measurements tend to occur in the latter part of the series, and the magnitude of the lowest annual maximum flows for the last 15 years seems to be higher than previous years.

As we have annual maximum data, we fit various stationary and non-stationary GEV distributions to the data. We assume

$$F_y(r) = P(M_y \leq r) = \exp\left\{ - \left[1 + \xi \left(\frac{r - \mu_y}{\sigma_y} \right) \right]^{-1/\xi} \right\}, \tag{8}$$

where μ_y and σ_y are (possibly time-varying) location and scale parameters repectively, ξ controls the tail behavior, and $\left[1 + \xi \left(\frac{x - \mu_y}{\sigma_y}\right)\right] > 0$. A positive ξ indicates a heavy tail. We restrict ourselves to the case where ξ is not time-varying, although the model could be generalized to allow this parameter to vary with time as well.

[1] http://nwis.waterdata.usgs.gov/nwis

[2] The real-time data is marked as "Provisional and subject to revision"

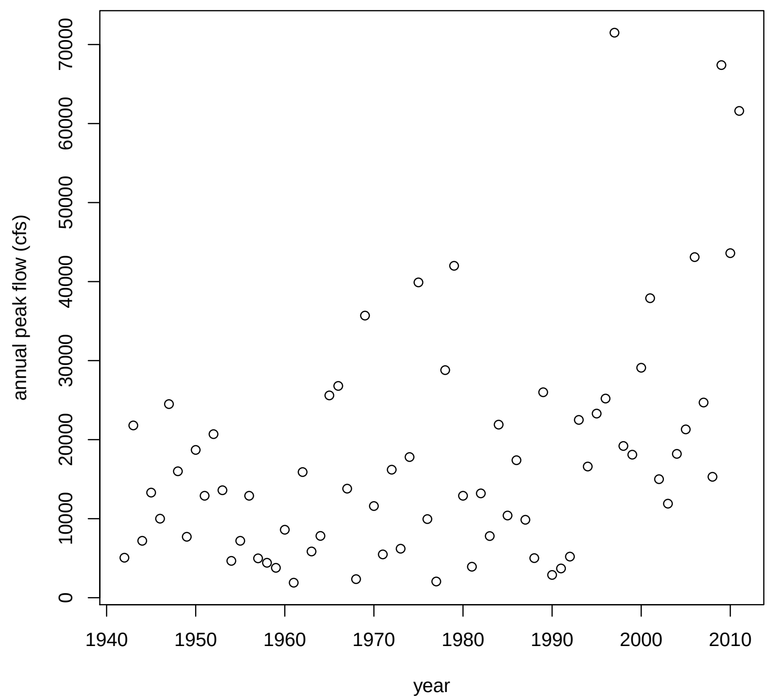

Figure 1: Annual peak flow measurements for the Red River at Halstad, MN.

3.1 The Stationary Model

We begin by fitting a stationary model to this annual maximum series. Via numerical maximum-likelihood, we fit the stationary GEV; that is, $\mu_y = \mu$ and $\sigma_y = \sigma$ in (8). Our estimates are $\hat{\mu} = 10392$, $\hat{\sigma} = 7924$ and $\hat{\xi} = 0.323$ with respective standard errors of 1095, 944, and 0.137. The log-likelihood of the fitted model is -751.59. We note that the MLE for ξ seems a bit higher than tail parameter estimates typically associated with river flow data. Also note the large uncertainty associated with this difficult-to-estimate parameter.

Given a return period m, plugging $\hat{\mu}$, $\hat{\sigma}$, and $\hat{\xi}$ into (8), setting equal to $1-1/m$ and solving yields the estimated m-year return level. We will focus on the 50-year return period, however, analogous analyses could be performed any return period. Our point estimate for the 50-year return level is 72306 cfs, and we note the largest observation in our 70-year record is 71500 cfs.

Approximate confidence intervals for the return level can be obtained by the delta method (Casella and Berger, 2002, Section 5.5.4) which relies on the asymptotic normality of maximum-likelihood estimators and produces a symmetric confidence interval. Alternatively, profile likelihood methods (Coles, 2001, Section 2.6.6) provide asymmetric confidence intervals, which better capture the skewness generally associated with return level estimates. For later comparison with the non-stationary cases we employ only the delta method procedure here, as profile likelihood methods for non-stationary measures of risk have not been developed. The delta method yields a 95% confidence interval of (33324, 111286).

3.2 A Nonstationary Model

We next fit a non-stationary model to the data. We allow the location parameter in (8) to be a linear function of time; that is, $\mu_y = \beta_0 + \beta_1(y - 1942)$, where y denotes year. Later we will consider other non-stationary models.

Again we fit the model using numerical maximum likelihood and use the maximum likelihood estimates from the stationary model and $\beta_1 = 0$ as the initial values for the numerical optimization procedure. Our estimates are $\hat{\beta}_0 = 7976, \hat{\beta}_1 = 106, \hat{\sigma} = 8630$, and $\hat{\xi} = 0.216$, with respective standard errors of 2247, 65, 1024, and 0.132. The log-likelihood of the model is -749.90. Model selection procedures such as the AIC (Akaike, 1974) would select the model with time-varying μ as the better model (AIC scores of 1509.2 and 1507.8 respectively for the stationary and time-varying μ models, lower is better). Interestingly, the standard error associated with β_1 would fail to reject a test with the null hypothesis that the slope is 0, as a Gaussian-based 95% confidence interval for this parameter estimate is (-21.4, 233.4).

3.2.1 Communicating Changing Risk

Next, we find the level for each year which has a 0.02 probability of exceedance (that is, a 1-in-50 chance) according to our fitted non-stationary model. Substituting our estimates for $\hat{\mu}_y = \hat{\beta}_0 + \hat{\beta}_1 y, \hat{\sigma}$, and $\hat{\xi}$ into (8) and setting equal to 1 - 0.02, we obtain

$$\hat{r}_{0.02}(y) = 106(y - 1942) + 60821.$$

The only peak flow to exceed the 0.02 probability-of-exceedance level is that of 1997 where the measurement of 71500 exceeds model's point estimate of $\hat{r}_{0.02}(1997) = 66651$. The 2009 peak flow of 67400 nearly exceeds $\hat{r}_{0.02}(2009) = 67923$.

It is interesting that $\hat{r}_{0.02}(y)$ does not exceed the stationary model's 50 year return level estimate of 72306 until the year 2050. The fact that the non-stationary model's estimated 0.02 probability-of-exceedance levels for the entire data record (1942-2011) are lower than the 50-year return level estimate of the stationary model can largely be attributed to the difference in the estimates of ξ.

Confidence intervals for the 0.02 probability-of-exceedance level for any particular year can be obtained via the delta method. The 95% confidence interval is (28989, 92652) for the year 1942, is (39183, 97148) for 2011, and extrapolating the model into the future yields (44996, 101978) for 2061.

3.2.2 Return Period as Expected Waiting Time

As explained in Section 2.2, design criteria may require that a return level be estimated for a given return period. In the non-stationary setting, any definition of return period has to be associated with a specific time. The expected waiting time definition of return period has a specific starting time, as it is an infinite sum beginning in a particular year. The expected number of events definition has a specific time interval corresponding the the years over which the sum is calculated.

We extrapolate the trend in our model and use the expected waiting time definition of return period to calculate the 50-year return level beginning in 2011. So that the equations in Section 2.2.1 make sense, we redefine y to be the number of years since 2011. We let $L = 200$, and get an estimated 50-year return level of 73150 cfs. That is, beginning in 2011, the expected waiting time until we see an exceedance of 73150 cfs is 50 years. With $L = 200$, the difference between the upper and lower bounds on our expected waiting time is less than 0.1 years.

Olsen et al. (1998) does not consider the uncertainty associated with the return level calculated from (3). We wish to employ the delta method to produce a confidence interval, but cannot do so directly since (3) does not yield an explicit expression for the m-year return level r_m. In the appendix, we show how the delta method can be used implicitly to obtain the variance of the return level. Using this method, we obtain a 95% confidence interval of (44383, 101916) for the 50-year return level beginning in 2011.

3.2.3 Return Period as Expected Number of Events

We now change our definition of return period in the non-stationary case to be the amount of time for which the expected number of events is one. Again extrapolating the trend in our model into the future, we calculate the return level associated with the 50-year return period from 2012-2061. Using (7), setting $m = 50$ and numerically solving yields $\hat{r}_m = 70950$ cfs. That is, in the 50-year period from 2012-2061, the expected number of exceedances of 70950 cfs is one.

Parey et al. (2010) employ a bootstrap method to obtain confidence intervals for return periods defined in this manner. We will instead resort to the implicit delta method outlined in the appendix. The advantage of the bootstrap approach is that it will likely result in an asymmetric confidence interval, better capturing the skewness

Model	$\hat{\mu}_y$	$\hat{\sigma}_y$	$\hat{\xi}$	$\ell(\boldsymbol{\theta};\boldsymbol{m})$	AIC
1	10392	7924	0.323	-751.59	1509.18
2	$7975 + 106(y - 1942)$	8630	0.216	-749.90	1507.80
3	9681	$4480 + 109(y - 1942)$	0.189	-748.70	1505.40
4	$8854 + 78(y - 1942)$	$4060 + 133(y - 1942)$	0.133	-746.53	1503.06

Table 1: Maximum likelihood parameter estimates, log-likelihood values, $\ell(\boldsymbol{\theta};\boldsymbol{m})$, and AIC values for the four models fit to the annual maximum series.

likely in the distribution of the return level estimate. The disadvantage of the bootstrap is that since we are dealing with only annual maximum data, our sample is rather small for bootstrapping. Bootstrapping heavy-tailed phenomena such as this river-flow data is particularly difficult, and Resnick (2007, Section 6.4) gives a thorough investigation. Our opinion is that any method for generating confidence intervals has drawbacks, but most methods provide the user a useful measure of uncertainty for the return level estimate. The implicit delta method yields a 95% confidence interval of (42458, 99441) for the 50-year return level associated with the period from 2012-2061.

3.3 Other Possible Non-stationary Models

The model in Section 3.2 is not the only way to model non-stationary behavior in the annual maximum time series. We consider two additional models. In the first, we hold the location parameter μ_y constant and allow the scale parameter σ_y to vary linearly with year. In the second, we allow both μ_y and σ_y to vary linearly with year. One could also consider other parametric forms for the behavior of these parameters or could consider non-parametric representations such as the work done by Chavez-Demoulin and Davison (2005). However, extrapolating non-parametric approaches to consider future risk could prove challenging. Additionally, one could consider allowing the tail parameter ξ to vary with time, which would imply that the fundamental behavior of the tail is changing. Cooley and Sain (2010) and Schliep et al. (2010) construct models where ξ is allowed to vary spatially, but not in time. Knowing that ξ is difficult to estimate even when held constant, we choose to allow only μ_y and σ_y to vary in time.

Table 1 gives the parameter estimates for these new models as well as the stationary and non-stationary models fit in Sections 3.1 and 3.2. All parameters estimated via numerical maximum likelihood. If the AIC is used as a strict model-selection criterion, each successive model is an improvement over the previous model. Figure 2 shows the QQ plot for each of the four models after transformation to standard Gumbel to account for non-stationarity (Katz, 2011) as well as the occurrence year of the largest five empirical quantiles (after transformation). That the largest empirical quantiles for the non-stationary model (Figure 2, upper-left) all occur after 1997 is some indication that the stationary model is failing to capture changing behavior, although the largest three quantiles for each of the non-stationary models are 1997, 2009, and 2011.

For each of the four models, we fit the 0.02 probability-of-exceedance level for the years 1942-2061. We also use the delta method to obtain 95% confidence intervals for these levels. Figure 3 shows the results for each of the four models. Clearly, the different non-stationary models give very different risk estimates over this period.

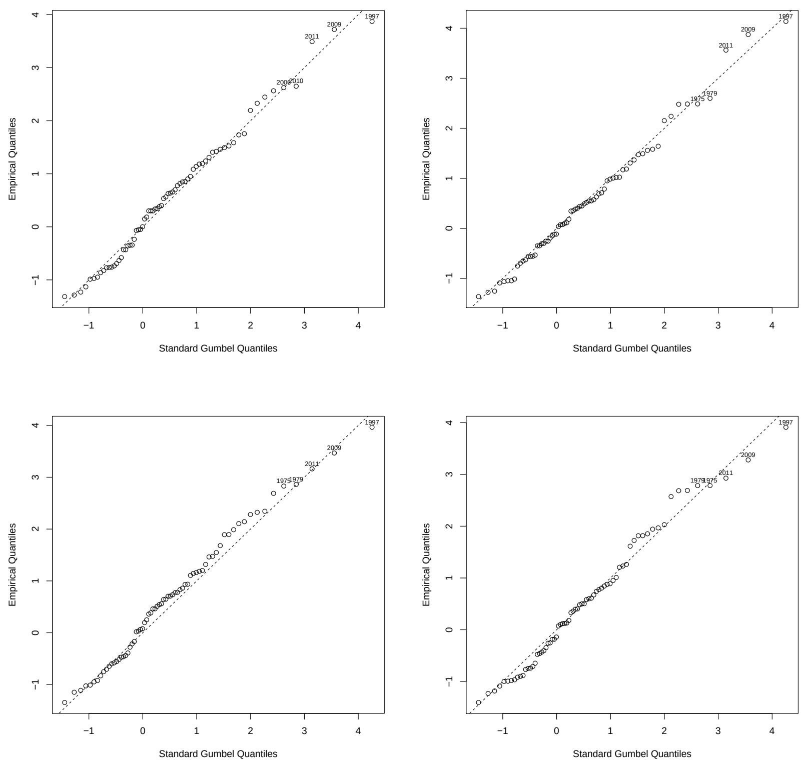

Figure 2: QQ plots based on transformation to a standard Gumbel distribution for the stationary model (top left), time varying μ model (top right), time-varying σ model (bottom left), and time-varying μ and σ model (bottom right). The occurrence year of the largest five empirical quantiles (after transformation) are also shown.

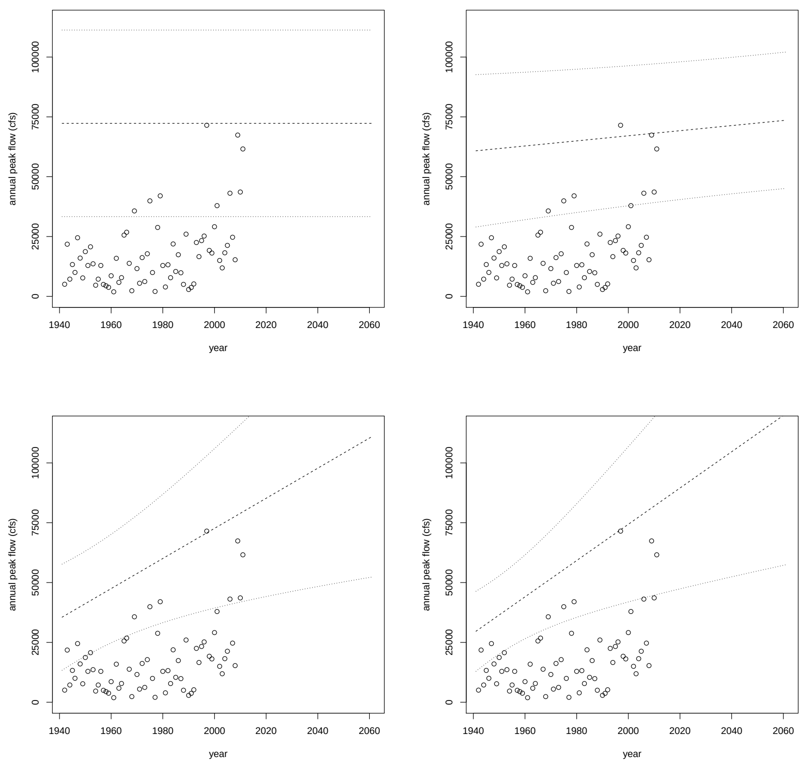

Figure 3: Yearly 0.02 probability-of-exceedance levels for the stationary model (top left), time varying μ model (top right), time-varying σ model (bottom left), and time-varying μ and σ model (bottom right).

4 Discussion

The aim of this chapter has been to investigate the concept of return period in the context of non-stationarity due to climate change. Our example is drawn from hydrology, but the ideas for conveying risk under non-stationarity can be used for most any application, for instance precipitation or temperature data. Also, the methods for conveying risk are not tied to fitting annual maximum data with a GEV distribution. Equations (3) and (7) require only estimates of $F_y(r_m)$, which could be obtained from fitting a GP or a log-Pearson type III distribution. Nor are the methods tied to the notion of return period being defined in terms of the annual maximum. Equations (3) and (7) could be adapted for use with the ARI rather than the AEP.

The Red River data example not only illustrates the different ways to communicate risk under non-stationarity, but also illustrates the questions that arise from fitting a non-stationary model and using it to convey future risk. A question that naturally arises regards extrapolating the fitted trend into the future. Figure 3 shows there is a significant effect on our risk measures based on which trend we choose. One of the basic lessons taught in an introductory regression course is the danger of extrapolating a model beyond the range of the data. Like it or not, estimating future risk requires extrapolation. All one can do is be clear in the assumptions made, and convey the uncertainty as best as one can.

Throughout, we have used a data-based approach. One can assess the fit of the four models by using tools such as the AIC, but we would be hesitant to select a model based *solely* on this or similar model selection criteria. One should always keep in mind that because extremes are rare, we are always data-poor when analyzing extremes, and thus it may be helpful to include outside expert knowledge into the model selection process.

The Red River example is interesting because, to our knowledge, the recent flood activity has not been directly linked to climate change. The Red River flows north, and floods most often occur in the spring when snow and ice begins melting in the south and the river remains frozen in the north. The recent floods do not seem to be caused by the exact same circumstances. The 1997 flood is largely attributed to snowmelt and extreme temperatures. The 2009 flood resulted from a combination of saturated and frozen ground, and snowmelt excerbated by rainfall. And the 2011 flood is at least partially blamed on high soil moisture in the previous year.

In part because we have used a data-driven approach, we have chosen our parameters to be linear functions of time. If extreme behavior is changing due to climate change, it is unlikely that they are behaving linearly. The recent mean temperature record, for example, does not appear to be increasing linearly but rather seems to have periods of rapid increase followed by plateaus. Still, linear trends may still be useful as models for long-term behavior and this is especially useful for extrapolation.

Another approach would be to use something other than time as the covariate in the model. For instance, one could imagine linking temperature data directly to CO_2 level rather than time. However, linking to a climatological covariate makes extrapolation into the future more difficult, as one would need to extrapolate the covariate as well. No obvious climatological covariate comes to mind for the Red River application.

There are other questions that arise with our particular application. The Red River was selected specifically because we were aware of recent flood activity. It is natural to ask if the non-stationarity we are seeing is due to selection bias and results from chance alone. Another question deals with the data itself. There is a well-documented flood in 1950, but the peak flow for this year is recorded as 18700 cfs, a rather unremarkable level. If this data point is in error and the 1950 level is in fact much higher, this would drastically impact our non-stationary model.

Finally, one can ask which method of communicating risk is best? The yearly probability-of-exceedance level communicates most clearly the changing nature of risk, but is less useful for design criteria. The two return period definitions each have their advantages. The expected waiting time definition is more closely linked to a lifespan calculation than the expected number of events definition. However, the expected waiting time definition extrapolates the trend indefinitely, whereas the expected number of events definition only extrapolates over the m years used in the calculation.

Acknowledgements: The author would like to thank Rick Katz, National Center for Atmospheric Research, for a helpful initial discussion, and for informing the author about important references. The author also thanks the editors for the invitation to write this chapter and an anonymous reviewer for helpful comments. The author acknowledges support from National Science Foundation grant DMS-0905315.

A Appendix

A.1 Expansion of (2)

Here we show how the expansion of (2) results in (3).

$$
\begin{aligned}
E[T] &= \sum_{t=1}^{\infty} t \prod_{y=1}^{t-1} F_y(r)(1 - F_t(r)) \\
&= 1(1 - F_1(r)) + 2F_1(r)(1 - F_2(r)) + 3F_1(r)F_2(r)(1 - F_3(r)) + \\
&\quad\; 4F_1(r)F_2(r)F_3(r)(1 - F_4(r)) \ldots \\
&= 1 - F_1(r) + 2F_1(r) - 2F_1(r)F_2(r) + 3F_1(r)F_2(r) - 3F_1(r)F_2(r)F_3(r) + \\
&\quad\; 4F_1(r)F_2(r)F_3(r) - 4F_1(r)F_2(r)F_3(r)F_4(r) + \ldots \\
&= 1 + F_1(r) + F_1(r)F_2(r) + F_1(r)F_2(r)F_3(r) + F_1(r)F_2(r)F_3(r)F_4(r) + \ldots \\
&= 1 + \sum_{i=1}^{\infty} \prod_{y=1}^{i} F_y(r).
\end{aligned}
$$

A.2 Implicit Delta Method

Let $\boldsymbol{\theta}$ be a d-dimensional parameter vector for a particular model, let $\hat{\boldsymbol{\theta}}$ be the maximum likelihood estimates, and let $V(\hat{\boldsymbol{\theta}})$ be the approximated covariance matrix of these estimates, obtained by inverting the hessian from the numerical optimization procedure.

The m-year return level is a function of the parameter vector: $r_m = f(\boldsymbol{\theta})$. Denote by ∇r_m the gradient of this function. The delta method says the variance of r_m is approximately

$$\nabla r_m^T V(\hat{\boldsymbol{\theta}}) \nabla r_m.$$

Unfortunately, we do not have explicit expressions for the function f for either of the m-year return level definitions in the non-stationary case. Rather, equations (3) and (7) can be generalized as $m = g(\boldsymbol{\theta}, r_m)$. Let m_0 be our desired return period and let r_{m_0} denote the corresponding return level. Knowing g we are able to find $\partial m/\partial r_m$, and $\partial m/\partial \theta_i$ for $i = 1, \ldots, d$. Thus

$$\begin{aligned}\left.\frac{\partial r}{\partial \theta_i}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, m=m_0} &= \left.\frac{\partial r}{\partial m}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, m=m_0} \left.\frac{\partial m}{\partial \theta_i}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, m=m_0} \\ &= \left(\left.\frac{\partial m}{\partial r}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, r=r_{m_0}}\right)^{-1} \left.\frac{\partial m}{\partial \theta_i}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, m=m_0},\end{aligned}$$

which allows us to calculate the needed gradient.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Balkema, A. and De Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5):792–804.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., Waal, D. D., and Ferro, C. (2004). *Statistics of Extremes: Theory and Applications.* Wiley, New York.

Bonnin, G., Todd, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D. (2004). *NOAA Atlas 14, Precipitation Frequency Atlas of the United States, Volume 1.* U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland.

Casella, G. and Berger, R. (2002). *Statistical Inference, 2nd Edition.* Duxbury, Pacific Grove, CA.

Chavez-Demoulin, V. and Davison, A. (2005). Generalized additive models for sample extremes. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 54(1):207–222.

Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics. Springer-Verlag London Ltd., London.

Cooley, D. and Sain, S. R. (2010). Spatial hierarchical modeling of precipiation extremes from a regional climate model. *Journal of Agricultural, Biological, and Envrionmental Statistics*, 15:381–402.

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

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the larges or smallest members of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180–190.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *The Annals of Mathematics*, 44(3):423–453.

IPCC (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller, editors.

Katz, R. (2011). Statistical methods for nonstationary extremes. In Sorooshian, S., Easterling, D., AghaKouchak, A., Schubert, S., and Hsu, K., editors, *Hydrologic Extremes in a Changing Climate*, XXX. XXX, XXX.

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

Mays, L. W. (2001). *Water Resources Engineering*. Wiley.

Olsen, J., Lambert, J. H., and Haimes, Y. (1998). Risk of extreme events under nonstationary conditions. *Risk Analysis*, 18:497–510.

Parey, S., Hoang, T. T. H., and Dacunha-Castelle, D. (2010). Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21:698–718.

Parey, S., Malek, F., Laurent, C., and Dacunha-Castelle, D. (2007). Trends and climate evolution: Statistical approach for very high temperatures in france. *Climatic change*, 81(3):331–352.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3:119–131.

Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

Schliep, E., Cooley, D., Sain, S. R., and Hoeting, J. A. (2010). A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling. *Extremes*, 13:219–239.

Smith, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, 4:367–393.

Vogel, R. and Wilson, I. (1996). Probability distribution of annual maximum, mean, and minimum streamflows in the united states. *Journal of Hydrologic Engineering*, 1(2):69–76.