1. List of authors:
   A. Behailu B. Bekera (Corresponding Author)
      University
      1776 G St., NW #101
      Washington, DC, USA, 20006
      Email: bttty@gwu.edu
      Phone: +1(202)-834-7710
      Fax: +1(202)-994-0245

   B. Royce A. Francis
      University
      1776 G St., NW #101, Washington, DC, USA 2006
      Email: seed@gwu.edu
      Phone: +1(202) 994-3892

   C. Olufemi A. Omitaomu (omitaomua@ornl.gov)
      Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak
      Ridge, TN 37831, USA.
      Department of Industrial and Systems Engineering, University of Tennessee, Knoxville,
      TN 37996, USA
      Phone: +1(865)-241-4310
      Fax: +1(865)-241-6261

Brief biographical notes about authors:

- Behailu B. Bekera is a doctoral student at Engineering Management and Systems
  Engineering Department of The George Washington University.
- Royce A. Francis, Ph.D., is an Assistant Professor at Engineering Management and
  Systems Engineering Department of The George Washington University.
- Olufemi A. Omitaomu, Ph.D., is a Senior Research Scientist at Computational Sciences
  and Engineering Division of the Oak Ridge National Laboratory.

Acknowledgements

The authors acknowledge the George Washington University School of Engineering and Applied
Science for the financial support of this work. This work was also partially supported by
National Science Foundation Award #1031046.
Drought Risk Modeling for Thermoelectric Power Plants Siting Using an Excess Over Threshold Approach

Abstract

Water availability is among the most important elements of thermoelectric power plant site selection and evaluation criteria. With increased variability and changes in hydrologic statistical stationarity, one concern is the increased occurrence of extreme drought events that may be attributable to climatic changes. As hydrological systems are altered, operators of thermoelectric power plants need to ensure a reliable supply of water for cooling and generation requirements. The effects of climate change are expected to influence hydrological systems at multiple scales, possibly leading to reduced efficiency of thermoelectric power plants. In this paper, we model drought characteristics from a thermoelectric systems operational and regulation perspective. A systematic approach to characterize a stream environment in relation to extreme drought occurrence, duration and deficit-volume is proposed and demonstrated. This approach can potentially enhance early stage decisions in identifying candidate sites for a thermoelectric power plant application and allow investigation and assessment of varying degrees of drought risk during more advanced stages of the siting process.

Keywords: Drought modeling, facility siting, nuclear power plants, thermoelectric systems, climate change, water resources, extreme events

1. Introduction

As concerns over climate change continue to grow, private and public operators of critical infrastructure systems seek to make their systems reliable and resilient in the face of unforeseen consequences. A critical scenario in this regard is a drought-driven energy crisis in countries where a considerably large portion of the electricity is produced by thermoelectric power plants (Badr et al. 2012). In the USA, for example, such power plants account for approximately 88% of the total energy production. This raises concerns as the power plants heavily depend on water for cooling purposes and water supply has already become a primary issue due to competing demands, forecast uncertainty and high costs of developing additional or upgrading existing infrastructures (Yang & Dziegielewski 2007). Energy production requirements are in competition with other water uses that may be stressed by climatic changes, population increases, and other economic demands, including but not limited to drinking water needs, industrial, agricultural and recreational purposes. Increased competition coupled with the effects of climate change is expected to influence hydrological systems at multiple scales, possibly leading to reduced efficiency of thermoelectric power plants (Feeley et al. 2008). Alleviating
such concerns demands proper understanding, modeling and analysis of drought with respect to thermoelectric power plants drought risks. This paper discusses methodological challenges in defining, characterizing and modeling drought aspects in the aforementioned context and is a critical building block in drought risk analysis for thermoelectric power plants planning and facility siting.

2. Drought as a System of Systems Problem

Proper functioning of a thermoelectric system depends on appropriate integration and interoperability of a set of interdependent systems including a reactor system, cooling system, aquatic system, environmental safety system, the climate system, and hydrological systems. The reactors for instance produce the heat required to turn water into steam which then drives the generators to produce electricity. The cooling system ensures the steam is condensed back to water while it avoids conditions that will make it approach maximum temperature threshold during operation. Sustainable cooling water supply on the other hand hinges on the normality of the climate system and parsimonious withdrawal by other users for irrigation, industry, domestic and public supply. In addition, nuclear facilities are subject to the National Environmental Policy and other appropriate local and state regulations. As such, energy companies are required to first allow used water to cool so that it will not affect life in the nearby aquatic system. The waste management system, tasked with storing and disposing of radioactive by-products, is part of the environmental safety system which in itself involves various stakeholders. Proper function of any of these systems relies on one or more of the rest and, as described above, all of them must appropriately be interconnected and integrated for a thermoelectric system to work. In short, suffice to say that a thermoelectric system certainly demonstrates the essential characteristics of a system of systems as widely discussed in the literature (Keating & Katina 2011; Haimes 2012).

Absence of adequate cooling water would force a thermoelectric power plant to cut back on production or even shutdown completely (Union of Concerned Scientists, 2011). Such a problem arises from reduction in the amount of water in the hydrological systems or simply because the water is so hot that it is not useful for cooling proposes. In recent years, a number of power plants such as the Millstone Nuclear Power Station in Connecticut, McGuire Nuclear Power Plant operated by Duke Energy, and four reactors operated by Progressive Energy Inc., have had to shut down or come close to doing so because of cooling water issues (Union of Concerned Scientists 2007). Whether it is caused by climatic factors leading to heating up of the water body or drying up of the stream environment or disruptions of the consumption patterns due to increasing demands by public, insufficient availability of cooling water debilitates normal operations of thermoelectric systems and therefore is a problem of a system of systems.

3. Background and Research Objective

Although detailed modeling of drought occurrence and its severity is extremely valuable, development of such models is challenging and computationally expensive. The vast majority of
existing hydrological models focus on estimating the most extreme drought scenarios (i.e., longest duration and largest deviation from normal water supply) in order to set design basis requirements for water storage systems such as reservoirs. As such, while significant progress has been made in understanding drought frequency and severity, the forecasting aspect, which is necessary for mitigation and adaptation planning, is yet to be adequately addressed (Panu & Sharma 2002). Besides, thermoelectric systems are often operated differently compared to the hydrological networks, which is one key reason for lack of operational applications of some of the existing hydrological models in the power utility industry. The former focuses on satisfying electricity demand whereas the latter monitors and studies short term water availability and future sustainability issues.

The objective of this paper is to elucidate the detailed assumptions risk analysts must make when evaluating drought risk to nuclear power plants through the construction of a model based on Excess Over Threshold (EOT) approach, also used in (Zelenhasić & Salvai 1987; Tallaksen et al. 1997; Cebrián & Abaurrea 2013). This approach, commonly used in hydrological networks studies, will be adapted and discussed in the context of thermoelectric power plants operation. The model proposed in this paper builds upon prior literature and is developed to provide an appropriate drought risk modeling framework from the perspective of thermoelectric power plants operation. The next sections are organized as follows: Section 4 presents a brief synopsis of drought definitions and its key components; Section 5 summarizes previously suggested modeling approaches and identifies the shortcomings this paper intends to alleviate; Section 6 presents our proposed modeling approach; and in Section 7, we validate the approach using daily streamflow series data obtained from the USGS surface water database. The last two sections provide a summary and discussion of lessons for early site selection processes and drought risk analysis of thermoelectric systems followed by a few concluding remarks.

4. Drought definition and characteristics

There is currently no universally accepted definition of drought. Perhaps the most pervasive approach has been to describe drought generally as a significant shortage of surface water supply over a certain areal extent for an extended period of time (Salas et al. 2005). The definitions become somewhat more specific when a particular type of drought is considered. According to information obtained from the U.S. Drought Mitigation Center and most of the literature, there are four types of drought: meteorological, agricultural, hydrological and socio-economic (Nalbantis & Tsakiris 2009). Meteorological drought is defined in terms of the degree of deviation from the average of a meteorological drought variable (e.g. precipitation) for an extended period of time over a certain area (Smakhtin & Hughes 2007). Agricultural drought is focused on soil moisture level and typically defined by comparing daily precipitation values to evapotranspiration rates (Sun et al. 2012). On the other hand, hydrological drought deals with changes in the components of the hydrological system such as streamflow, groundwater and reservoir levels. Finally, socio-economic drought is defined as an imbalance between supply and
demand of goods as a result of a weather-related shortfall in water supply (Wilhite & Buchanan-Smith 2005).

In this paper, we employ the following statistical definition of drought. Drought is a complex random process characterized by its occurrence frequency, duration, deficit-volume, intensity and areal extent (Salas et al. 2005; Sharma 1997; Cancelliere et al. 2007). Frequency clearly refers to the number of occurrences over a given period of time. Deficit-volume, measures the degree to which a drought variable of interest (e.g. rainfall, streamflow) departs from a given lower-end threshold during a certain drought period. Drought intensity, usually given by the ratio of deficit-volume and duration, combines the effects of deficient quantity and length of time for which the drought persisted. It therefore provides a robust approach to compare severity of two or more different drought events.

5. Hydrological modeling approaches

Modeling drought is challenging for a number of reasons. First, drought is a result of interactions among many complex climatic and geographic elements. Second, drought takes time to develop and as such it is difficult to pinpoint distinct time of occurrence (Cebrián & Abaurrea 2013). Similarly, a drought period may, for a short period of time, be interrupted by wet days which could potentially be misleading and also makes it tough to discern the actual drought termination time. Consequently, an extended drought may be divided into a number of smaller, mutually dependent droughts. As a result, capturing the duration and deficit-volume pertaining to a drought event is not as easy as it sounds.

Nonetheless, several statistical methods of modeling drought events and its behaviors have been suggested over the past few decades. Each has added a significant contribution to drought study and brought to light various ways of modeling its specific characteristics or large droughts in general. Yevjevich (1967) and Sen (1976) suggested the theory of runs. According to this theory a drought period consists of a set of time series of a drought variable below a predefined threshold or truncation level. Regression-based methods have also been applied to relate drought variables to climatic and other factors (Mimikou et al. 1993). Kendall & J. A. Dracup (1992) proposed an alternating renewal-reward process that rotates between drought and high streamflow event durations. This model assumes that droughts and high streamflow events exhibit distinct properties and as such are governed by different probability laws. Chung & Salas (2000) also investigated severity of drought by developing a drought intensity function assuming time-dependent Poisson behavior of drought series. The Excess Over Threshold (EOT) method is yet another predominant approach in the literature (Zelenhasić & Salvai 1987; Tallaksen et al. 1997; Cebrián & Abaurrea 2013). Extreme Value Theorem (EVT) is the theoretical foundation behind EOT based on statistical models representing rare events that have heavy consequences (Davison & Smith 1990).
In EOT model, drought is defined as a cluster of dependent dry spells and is considered as a multivariate random event characterized by its duration, maximum intensity and corresponding deficit-volume. For a given threshold, \( r \), there exists a conditional distribution function, \( F_r(y) \) that represents excesses (that is, below threshold in case of drought modeling) of an iid random variable, \( X(t) \) (Këllezi & Gilli 2000).

\[
Y(t) = r - X(t)
\]  

(1)

\[
F_r(y) = P(r - X(t) \leq y(t) \mid X(t) < r), \quad 0 \leq y(t) \leq r
\]

(2)

EVT suggests that exceedances over a very extreme threshold, \( r \), are distributed following a Generalized Pareto Distribution (GPD) (Këllezi & Gilli 2000; Smith 1986).

\[
F_r(y(t)) \approx G_{\xi, \sigma, y(t)}, \quad r \to \infty
\]

(3)

\[
G_{\xi, \sigma, y(t)} = \begin{cases} 
1 & \text{if } \neq 0 \\
\left(1 + \frac{y}{\sigma}ight)^{-\frac{1}{\xi}} & \text{if } = 0
\end{cases}
\]

(4)

where \( \sigma \) and \( \xi \) represent the scale and shape parameters of the GPD respectively for \( 0 \leq y(t) \leq r \).

Moreover, according to this theory, the number of occurrence of these exceedances over extreme threshold fits well with a Homogenous Poisson Process (HPP). However, we know that a drought would constitute a number of dry spells (exceedances) and the Poisson behavior may be lost in the process of clustering several dependent dry spells into a single drought event. In other words, one may be concerned that occurrence of droughts (that is, collection of dry spells), may not be a point process and hence may not fit well with a Poisson process. As others (Zelenhasić & Salvai 1987; Cebrián & Abaurrea 2013; Sun et al. 2012; Smith 1986) have also argued that, if the Poisson process does no longer hold, it can be recovered by introducing a more extreme threshold. This is because, as the threshold becomes more extreme, there will be fewer total observations below such a limit and consequently fewer dry spells constitute a drought event. Eventually, however, if the threshold is extreme enough, the number of exceedances (dry spells) in any cluster converges to 0 or 1 and the clusters themselves become a single dry spell.

Among the advantages of application of modeling drought this way are the estimations of expected number of droughts in a given time length, approximations of expected recurrence time, expected duration, average water deficiency volume and intensity level for a certain degree drought. Estimation of drought intensity depends on how deficit-volume and drought duration are modeled. Typically, and as is the case with existing EOT approaches, one or more empirical truncation levels are introduced to determine onset and termination of such an event. Drought
length is approximated by the time between beginning and end of the event. This approach is inconsistent with drought duration from power plants operational point of view. Power plants are likely to shut down during drought conditions as water levels are expected to fall below intake depths from the surface (Badr et al. 2012). The inconsistency is because a certain drought period may temporarily be interrupted by normal water supply levels and as a result, the power plants may be turned on and off a number of times during large and extended drought period. For that reason, drought duration, from the point of view of power plants operation, is not equivalent to the time between drought onset and termination. For example, the Marked Cluster Poisson Model (Cebrián & Abaurrea 2013) uses an additional threshold to mark return of normal behavior (for example, 30% quantile of a time series of drought variable). In other words, it is used to determine the time a certain drought period comes to an end, thus affecting the length of and intensity associated with the drought. Consequently, such an approach to capture drought duration and resultant estimation of the intensity is misleading and inefficient from power plants operational perspective.

6. Proposed approach for modeling drought risk to thermoelectric systems

Our approach, also based on Excess Over Threshold method, aims to alleviate the aforementioned issues of duration modeling, onset and termination detection, and threshold selection. This is done by enabling drought duration, and hence intensity modeling in a manner consistent with regular operations of power plants. In addition, an intuitive and important persistency test is suggested on top of an empirical threshold to determine existence of sufficient signals about any occurrence of drought at a given point in time. The suggested approach in the following section consists of drought threshold selection, dry spells pooling mechanism, drought occurrence modeling, and drought duration and deficit-volume approximation techniques.

6.1 Empirical Threshold Selection

Many recent drought studies have applied threshold methods because of their advantage in using data more efficiently (Këllezi & Gilli 2000). Thresholds are critically important in drought modeling as the severity drought to be investigated is essentially determined by the selected truncation level. The mean or median (John A. Dracup et al. 1980) and fraction of the mean (Clausen & Pearson 1995) over a long period of time have been used before for truncation levels in drought analysis. These methods are hardly flexible for investigating varying degrees of droughts and, for that reason, are no longer preferred approaches. Another more commonly applied technique is taking the lower Q% quantile of a long-term time series data (Panu & Sharma 2002). So, the cutoff level would be selected such that only Q% of the entire data points are less or equal to it; but even more extreme quantiles may be necessary in order for the EVT based rules to hold. This is also our method of choice to define the cutoff level in modeling drought pertaining to power plants. Not only does this method provide the flexibility to investigate different droughts but also allows for efficient use of historical data for short-term and long-term drought mitigation and recovery planning.
Alternatively, such thresholds could also be obtained from the water flow requirements set by appropriate government agency for each and every power plant at the time their licenses were approved. However, most of these licenses are, for example, over 30 years old for nuclear power plants and most of the assumptions of stationarity have been repeatedly challenged by various water supply factors and issues discussed earlier on. Therefore, defining drought on the basis of such decades old design-basis requirements would be so unrealistic that it would be difficult to conduct a meaningful analysis (Milly et al. 2007).

6.2 Drought occurrence

As said before, the process by which drought events occur tends to be Poisson distributed. The rate parameter of this distribution will inform us about expected number of droughts per unit time. Among the well-known properties of Homogenous Poisson Processes (HPPs) is the superposition property--if there are \( n \) independent drought occurrences each with their own rate of arrival, the overall arrival processes will be a HPP with the sum of individual rates being its new rate of arrival. This property provides a powerful method for aggregating stream segment level information, perhaps with an assumption of independence, into system wide drought information. The authors are currently investigating this research related to ongoing drought risk analysis and infrastructure facility siting study with specific focus on power plants.

6.3 Pooling mechanism

Most of the challenges in EOT based drought modeling culminate in problems pertaining to clustering of dry spells. Some studies have suggested the use of an uninterrupted sequence of deficits to constitute a drought event (Sharma 1997). This would, however, divide up large drought area into several small drought areas and may offer wrong information about frequency of occurrence. This problem, however, is alleviated by introducing inter-dry spell time limit such that protracted droughts are not misconstrued into representing separate drought periods.

![Figure 1](image.png)

Figure 1. Pooling dry-spells to form a drought event and associated deficit-volume for segments of drought duration. The deficit-volume pertaining to the drought is not the entire shaded area because the deficit at a given point in time itself is reflective of previous shortages. We take the point where the deviation from the threshold streamflow is the highest as indicative of the deficit-volume for the drought event.
Our method of implementing this approach is as follows: First, we determine onset of drought when a streamflow series falls short of the threshold, \( X(p) \leq r, p \in T \) and when there are sufficient indications about its persistence. Drought develops over time and as such, there must be some indication of sustained concern about shortage of water supply and significant indication as to the occurrence of drought. For a very extreme, threshold one exceedance by itself can be of a huge concern. But we believe that if such a series drought occurs it must persist for some time. Hence, for a dry day (that is, \( X(p) \leq r \)) drought is said to occur on this day if it passes a simple yet important persistency test (that is, at least 4 of the first 30 days), including the first one, fall below the threshold as well. This criterion allows us to emphasize the evolving nature of drought, and to ensure that there is a reasonable evidence for the occurrence of drought to be attributed to a given dry day. In case this is not satisfied, an algorithm (See Appendix B) developed this way would simply move on to the following day and does the same persistency test until the most appropriate onset time is determined. Second, since a certain drought event may occasionally involve normal streamflows without necessarily indicating termination of a dry season, there needs to be some inter-event time limit such that two dry spells must always remain in the same dry season. The idea here is to avoid dividing up of dry spells that are likely to be attributed to a single drought event into two or more separate droughts. Due to an extreme threshold selection, the droughts considered are likely to recur in a number of years. For that reason, we assume six months of inter-event time limit (that is, dry spells within this limit must belong to the same drought period) are sufficient or, alternatively, the Ferro & Segers (2003) procedures can be used to obtain such a limit automatically. This keeps dry days within a short time span in the same drought period. Past this limit, the drought comes to an end when \( X(q) > r, q \in T, q > p \). From this point on, the next pooling ensues when \( X(s) < r, s \in T, s \geq q \) granted passing of the persistency test.

On the other hand, one may be concerned about \( X(t) \) being just slightly above the threshold for a prolonged time without solid justification of return of normal behavior. If this slight exceedance remains for a long time we assume it’s an indication of improving conditions. However, if it is temporary and the flow persistently drops back below the critical threshold it would be captured as the beginning of a new round of pooling.

5.4 Drought duration and deficit-volume modeling

In case of an unlikely but catastrophic drought event, power plants are expected to shut down completely as the water levels fall far below the normal intake depths from the surface (Badr et al. 2012). Hence, drought duration associated with the shutdown period, \( L(n) \), of the \( n^{th} \) drought event, is approximated by the total number of days below the threshold. Note that this approach of modeling drought duration is different from the typical approach that takes the difference between beginning and ending times of an event.
The volume of deficient water naturally changes as a drought event unfolds and such events are often identified with the point where the worst scenario takes place. With that notion of water resource managers in mind, the point where drought intensity is the highest is used to obtain the deficit-volume representative of the event.

7. Model validation and application

Streamflow (in volume per unit time) is generally considered the most significant drought variable to convey water quantity information (Nalbantis & Tsakiris 2009). The proposed methodology is applied on a daily streamflow series obtained from three USGS water gauges on the Tennessee River basin: 1) USGS Tennessee river at Whitesburg Alabama; 2) USGS Tennessee river at Chattanooga Tennessee; and 3) USGS Tennessee river at Savannah Tennessee. These locations are of particular interest as they correspond to the nearest USGS water gauge power plants operated by the Tennessee Valley Authority (TVA). The application was performed on a 72-year streamflow record for Whitesburg, 109-year record for Chattanooga and 77-year flow record for Savannah water gauges.

7.1 Drought events as Poisson process

The first task in the application of the model to the occurrence process is to identify the appropriate extreme threshold at which the occurrences are well approximated by a HPP. For Whitesburg, the first trials at 10\(^{th}\) percentile (424.75 m\(^3\)/s) and 5\(^{th}\) percentile (308.65 m\(^3\)/s) thresholds did not produce Poisson distributed droughts. In Figures 2(a) and 2(b), the probability plots of fitting exponential distribution to time between droughts show large divergence from the straight line, thus failing to qualify the drought occurrences as HPP. Therefore, more extreme threshold values were needed. We can observe from Figure 2(c) that, for \(r = 1.4^{th}\) percentile (196.2 m\(^3\)/s), the probability yields, despite a couple of outliers, an acceptable exponential distribution fit with the P-value of 0.16. At this threshold, a total of 20 drought occurrences are discovered over 72-years. In the other two cases, the best fits were obtained at 1.1\(^{th}\) percentile (199.69 m\(^3\)/s) for Chattanooga and at the 0.08\(^{th}\) percentile (240.51 m\(^3\)/s) for Savannah, TN (See Appendix A).

Next, we present Poisson process fits to the occurrences as demonstrated in Figure 3. Visually, we observe that the straight line, that is, an indication of constant rate of occurrence, is a reasonable approximation for mean arrivals of drought at some time, \(t\). Constant rate of occurrence in turn is indicative of homogeneity. Results for HPP vs. Non-HPP investigations are presented in Table 1. None of these locations indicate sign of non-homogeneity as a Chi-squared test for Power law and Laplace Test for Log-linear rates of occurrences were conducted at 5% significance level. The conclusions were the same at 10% significance level for Whitesburg and Chattanooga whereas Savannah rate of drought occurrence exhibited a Log-linear behavior. But the probability plot of the exponential distribution to Savannah time between droughts (See
Appendix A), at r=0.08th percentile, produced a very high P-value (0.938) that we preferred to proceed with homogenous assumption.
Figure 2. Whitesburg probability plots of exponential distribution (95% confidence interval) fitted to time between droughts data obtained at different threshold levels. Figures (a), (b) and (c) are obtained at threshold values of 10th, 5th and 1.4th percentiles respectively.
Figure 3. Mean arrival rate fit to drought occurrence process. The red line indicates the mean fit, while the blue line indicates the observed records. The y-axis indicates the number of occurrences, while the x-axis indicates the number of days observed since the beginning of the streamflow record.
Table 1. Results of model hypothesis testing. The null hypothesis ($H_0$), droughts occur according to HPP, and two alternative cases of non-HPP are tested. HPP is favored in all cases at 5% significance level. That means, at 5% significance level, there is no sufficient evidence to reject homogeneity in all three cases.

<table>
<thead>
<tr>
<th>Location</th>
<th>Model Comparison</th>
<th>Test Statistic Value</th>
<th>Preferred Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(all compared to Homogeneous Poisson)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chattanooga</td>
<td>Log-Linear</td>
<td>-0.32</td>
<td>HPP</td>
</tr>
<tr>
<td></td>
<td>Power Law</td>
<td>45.68</td>
<td>HPP</td>
</tr>
<tr>
<td>Savannah</td>
<td>Log-Linear</td>
<td>1.92</td>
<td>HPP</td>
</tr>
<tr>
<td></td>
<td>Power Law</td>
<td>22.17</td>
<td>HPP</td>
</tr>
<tr>
<td>Whitesburg</td>
<td>Log-Linear</td>
<td>0.41</td>
<td>HPP</td>
</tr>
<tr>
<td></td>
<td>Power Law</td>
<td>38.70</td>
<td>HPP</td>
</tr>
</tbody>
</table>

7.2 Mean time between drought events and expected number of drought events

The Maximum Likelihood Method of fitting a HPP to a stochastic process is applied to estimate the only model parameter, rate of arrivals ($\lambda$). The average recurrence time is obtained as the reciprocal of the mean arrival rate and expected number of arrivals of drought events, for a given time length, $t$, is given by $\lambda t$.

Table 2. Mean time between arrivals and expected number of drought occurrences

<table>
<thead>
<tr>
<th></th>
<th>Whitesburg (Threshold at 1.4th percentile)</th>
<th>Savannah (Threshold at 0.08th percentile)</th>
<th>Chattanooga (Threshold at 1.1th percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$ (# of arrivals per year)</td>
<td>0.2850</td>
<td>0.165</td>
<td>0.1857</td>
</tr>
<tr>
<td>Mean arrival time (Years)</td>
<td>3.51</td>
<td>6.04</td>
<td>5.38</td>
</tr>
<tr>
<td>Expected # of droughts in 50 years</td>
<td>14.25</td>
<td>8.25</td>
<td>9.29</td>
</tr>
<tr>
<td>= $\lambda t$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It can be seen from the results in Table-2 that Savannah is modeled with the most extreme threshold (hence more severe drought) and expects the fewest (8.25) droughts in 50 years. Whitesburg may expect drought (below 1.4th percentile threshold) every 3.51 years on average and a total number of 14.25 such droughts in 50 years. Chattanooga, which is where the TVA collects flow information to obtain best estimates for system wide water supply level, expects somewhat in the middle recurrence time (5.38) and expected number of droughts in 50 years (9.29).
7.3 Drought length

Clearly, time between droughts is not the same as drought length. We discussed the former in the previous sections, but one would be equally interested in how long a particular drought is expected to last. Several probability distributions were fitted to each set of duration data obtained after feeding a daily streamflow series data to the software program developed for this study. A Kolmogorov-Smirnov goodness of fit test reveal Inverse Gaussian, Log-logistic and Log-normal distributions to be the best candidates for generalizing the probability law governing durations of various droughts.

Table 3. K-S statistic for drought duration goodness of fit test using @Risk analysis tool. The top three distributions with very close K-S statistic are shown for each location.

<table>
<thead>
<tr>
<th>Location</th>
<th>Inverse Gaussian</th>
<th>Log-logistic</th>
<th>Lognormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whitesburg</td>
<td>0.183</td>
<td>0.126</td>
<td>0.16</td>
</tr>
<tr>
<td>Chattanooga</td>
<td>0.116</td>
<td>0.104</td>
<td>0.113</td>
</tr>
<tr>
<td>Savannah</td>
<td>0.140</td>
<td>0.123</td>
<td>0.144</td>
</tr>
</tbody>
</table>

The fit with Log-normal distribution, shown below in Figure 4, demonstrates both visually and statistically the significance of using the distribution to generalize the probabilities of drought durations. On top of that, Table-4 presents a 95% prediction interval for the drought durations at each location. This is computed first by transforming the duration data such that it is normally distributed. Second, a 95% confidence interval for the mean of the transformed data is calculated and the boundaries of this interval are the transformed back to the form of the original duration data so as to provide the 95% prediction interval.

The expected values of durations are 13.22, 20.22, and 17.51 days for Whitesburg, Chattanooga and Savannah drought durations respectively. Once again, the prediction interval at Chattanooga seems to be somewhat comprehensive in a sense that it, to a large degree, covers the prediction intervals for the other two locations. That is again consistent with the claim that Chattanooga provides system wide water supply information.
Figure 4. Chattanooga drought durations with a Log-normal distribution fit. The probability plot on the left resulted in a significantly high P-value (0.503) thereby qualifying the Log-normal distribution as a suitable candidate to generalize probabilities of drought durations.

Table 4. A 95% prediction interval for drought duration at each location

<table>
<thead>
<tr>
<th>Location</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Expected Drought Duration(Days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whitesburg</td>
<td>7.33</td>
<td>14.31</td>
<td>13.22</td>
</tr>
<tr>
<td>Chattanooga</td>
<td>8.64</td>
<td>20.38</td>
<td>20.2</td>
</tr>
<tr>
<td>Savannah</td>
<td>6.62</td>
<td>19.36</td>
<td>16.17</td>
</tr>
</tbody>
</table>

7.4 Drought deficit-volume

As expected, the GPD fits to the deficit-volumes with P-values, 0.837, 0.632, and 0.558 for Whitesburg, Chattanooga and Savannah respectively. These fits demonstrated in Figure 5 highlight the flexible nature of the GPD attributable to its shape parameter. Aside from describing different aspects of a drought, these distributions can be used to determine the return value of deficit-volume, duration or intensity for a given time length. The exceedance probability of a given distribution of a drought variable is given by \( P(X > x) = 1 - F(x) \). For a time length of \( t \), the expected number of droughts is \( \lambda t \). The return value we are interested in (that is, the duration or deficit that is exceeded at least once on average in the given time period) occurs if there is at least one drought occurrence during the specified time length. In other words, we are looking for the \( X \) value such that

\[
\lambda t \left( 1 - F(x) \right) = 1
\]  
(5)
So, this value of $X$ at given time, $t$, is obtained as follows.

$$X_t = F^{-1}(1 - \frac{1}{\lambda t})$$  \hspace{1cm} (6)

This is important because once the $\lambda$ parameter is estimated from the occurrence process, we easily can make statements about expected durations, deficit-volumes, and intensity levels expected to be exceeded at least once in a given time period.
8. Summary and Discussion

In this paper, we have defined drought duration in terms of the amount of time during which a thermoelectric power plant remains offline due to insufficient fresh water supply. This definition takes into consideration that in case of a protracted drought, these plants may be turned on and off several times before the dry season is permanently over. Our proposed modeling approach is in accordance with this insight and differs from the dominant approach that uses the entire time between drought onset and termination as drought duration. We believe that the proposed definition and subsequent modeling of drought is consistent, especially in extreme drought situations, with the operations of and risk management strategies related to thermoelectric power systems.

We also suggest that water flow requirements be revised regularly in accordance with changing needs for and increasing competition of water resources. The alternative duration and intensity modeling we have provided should provide the foundation for devising an empirical flow limit that power plant operators should maintain at all times. For instance, the return value equation can be used to identify a deficit-volume or drought duration corresponding to a given recurrence time. This information, given we know how much water is needed for a power plant to operate at full efficiency, can then be transformed to provide a reasonable and site specific flow requirements to maintained.

The number of drought occurrences is also modeled using the Homogenous Poisson Process whose unique superposition property could potentially allow conducting drought risk analysis at a larger scale. Moreover, this approach is intended to assist in identification, approval, and study of thermoelectric infrastructure facility sites. If an additional power plant is being planned, the
method proposed here can be used to evaluate potential sites against cooling water supply requirements. Areas with a tendency to experience inadequate water supply can be precluded at early stages of such projects and before a large commitment of resources is made. A regulatory body of government may also use these techniques to support licensing decisions.

9. Conclusion

Water availability is among the most important elements of thermoelectric power plant site selection and evaluation criteria. Candidate sites are often identified through a rigorous selection process that involves exclusion, and avoidance of non-suitable areas based on predefined selection and evaluation criteria. In many cases, reliable supply of cooling water is assumed if the power plant is in close proximity to large water sources such as a lake or an ocean. In this paper, we propose and demonstrate an alternative systematic approach to characterizing extreme drought occurrence in a stream environment incorporating duration and deficit-volume. This approach makes two contributions to thermoelectric systems related risk analysis. First, it enhances early stage decisions in identifying candidate sites for a thermoelectric power plant application. Second, it improves risk characterization during more advanced stages of the site selection process by enabling planners to investigate the implications of droughts of varying severity and frequency. Aside from these contributions, the authors recognize that more work needs to be done to systematically predict water availability and make siting decisions more robust. While these results are very encouraging, we are currently investigating if this approach can be generalized to other areas; if not, we are investigating why that may be the case. Furthermore, we are also investigating the implication of these results for existing water policies for the present and future power plants with respect to climate change. Uncertainty is a big issue with streamflow sensor data and hence we are also investigating how sensor data uncertainty could impact the overall results of our proposed models. Finally, the authors are also interested in the application of these results for future power plant facilities siting studies.
Appendix A

**Chattanooga:** More of the time-between drought occurrences data points converge towards the boundary lines as more extreme droughts are considered. The first two trials at the 10\textsuperscript{th} and 5\textsuperscript{th} percentile show that these types of droughts can’t be modeled as point processes as their time between arrivals are not exponentially distributed. A reasonable fit, as evidenced by a high P-value of 0.319 to the exponential distribution was attained at the streamflow value that is exceeded by 98.9% of all the data points (that is where \( r = 1.1 \textsuperscript{th} \) percentile).
Figure 6. Chattanooga probability plots of exponential distribution fitted to time between drought occurrence data obtained at different threshold levels. Figures (a), (b), (c) correspond to a threshold of 10th, 5th and 1.1th percentiles respectively.

![Exponential - 95% Confidence Interval](image)

Figure 7. Savannah probability plots of exponential distribution fitted to time between drought occurrence data obtained at different threshold levels. Figure 7 (a) is produced at the 3rd percentile value of $r$ whereas Figure 7 (b) indicates a very high P-value that is obtained at $r = 0.08^{th}$ percentile, 240.51$m^{3}$/s. Twelve droughts were discovered over 72 years. Occurrence of these can very well be approximated by a Poisson process.

Appendix B

In order to support the modeling process, we developed a simple yet important algorithm that takes a long term daily streamflow data and performs the following tasks to capture drought occurrence times. We present the general idea here so as to help in the process of reproducing the process: 1) designate a binary value for each day to indicate if it is below or above threshold (1 for below and 0 for above); 2) locate the first dry day and performs persistency check in the next 30 days; 3) mark beginning of drought if detected dryness is persistent enough or moves on to the following day to look for more appropriate onset date; 4) search for an exit signal that is any day with above threshold streamflow at least 6 months after the onset was detected; 5) once the current drought period comes to an end, the algorithm is repeated until the following drought occurrence dates are identified.
Figure 8. Flow chart showing how long-term historical data can be examined for drought indications and occurrences. A streamflow series is first read in to the program. If a streamflow value is below the threshold, $R$, persistency test is performed. Passing this test confirms existence of sufficient evidence that drought occurred. This then invokes the inter-dry spell function that enforces the 6 months rule. Failing the test (i.e., no sufficient evidence to declare occurrence of drought on that particular day) takes the process back to the start where the next streamflow is read.
References


