Uncertainties in estimating design droughts

Devanmini Halwatura
Centre for Water in the Minerals Industry, Sustainable Mineral Institute, The University of Queensland, d.halwatura@uq.edu.au

Alex M. Lechner
Centre for Social Responsibility in Mining, Sustainable Mineral Institute, The University of Queensland, a.lechner@uq.edu.au

Neil McIntyre
Centre for Water in the Minerals Industry, Sustainable Mineral Institute, The University of Queensland, n.mcintyre@uq.edu.au

Sven Arnold
Centre for Water in the Minerals Industry, Sustainable Mineral Institute, The University of Queensland, s.arnold@uq.edu.au

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Part of the Civil Engineering Commons, Data Storage Systems Commons, Environmental Engineering Commons, Hydraulic Engineering Commons, and the Other Civil and Environmental Engineering Commons

Halwatura, Devanmini; Lechner, Alex M.; McIntyre, Neil; and Arnold, Sven, "Uncertainties in estimating design droughts" (2016). International Congress on Environmental Modelling and Software. 23.
https://scholarsarchive.byu.edu/iemssconference/2016/Stream-B/23

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Uncertainties in estimating design droughts

Devanmini Halwatura\(^a\), Alex M. Lechner\(^b\), Neil McIntyre\(^c\), Sven Arnold\(^a\)

\(^a\) Centre for Water in the Minerals Industry, Sustainable Mineral Institute, The University of Queensland, Australia (d.halwatura@uq.edu.au, n.mcintyre@uq.edu.au, s.arnold@uq.edu.au)

\(^b\) Centre for Social Responsibility in Mining, Sustainable Mineral Institute, The University of Queensland, Australia (a.lechner@uq.edu.au)

Abstract: Severity-duration–frequency (SDF) curves of periods of rainfall deficits are useful tools for drought analyses. However, accuracy of these curves are affected by uncertainties associated with a number of factors including: (i) the choice of drought index, (ii) the sampling error due to the limited length of observation data, (iii) the effects of aggregation of data with respect to drought timescales, (iv) the selection of marginal probability distribution functions of drought severity and duration, (v) and the type of copula used to approximate the dependency between severity and duration. In this paper we assess the impact of these uncertainties on estimates of Recurrence Intervals (RI) of drought events by comparing three drought indices (the Standardised Precipitation Index, Reconnaissance Drought Index, and Standardised Precipitation Evaporation Index), three timescales (three, six and twelve months), four marginal probability distribution functions (extreme value, logistic, gamma, lognormal), and two types of copulas (Gumbel and Frank). We assessed all parameterization combinations in relation to their resulting drought RIs of mild, moderate and extreme drought events for 11 sites across Eastern Australia and compared them with a selected baseline value. Across the three selected drought indices there was no difference between the RI derived with Gumbel and Frank copulas for all the sites. For mild drought events, there was no difference between the RIs derived with two drought indices (SPI and SPEI) and distribution functions with the baseline case, whereas SDF curves showed the highest uncertainty with the 12 month time scale. Design droughts or SDF curves are a critical part of any drought analysis for the management of natural systems in regions that can potentially experience water-limited conditions. Explicitly considering the uncertainties involved in developing design droughts is important when assessing the risk of ecosystem failure due to drought events.

Keywords: design droughts; severity duration frequency (SDF) curves; uncertainty analysis; drought index

1 INTRODUCTION

Droughts are different from other natural hazards because they often exhibit prolonged gradual build-ups of deficits in rainfall. These rainfall deficits affect water supply and may result in severe and long-term financial hardship for farmers and impact natural environments (Passiouara, 2007). In recent years almost all the continents have been affected by severe droughts which have caused billions of dollars of damage annually; and it is expected that the severity and frequency of droughts are going to increase in many regions of the world in the future (Heim, 2002). Therefore the topic of drought analysis has seen increased interest in recent years.

There are four rationales for drought analyses (Byun and Wilhite, 1999): (1) to understand the atmospheric circulation associated with drought occurrences, (2) to understand the frequency and severity of drought, (3) to identify impacts of drought, and (4) to reduce the impacts through preparing for droughts and mitigating droughts. Most drought related studies focus on analysing frequencies of droughts of a given severity and duration, and identifying possible drought impacts (Byun and Wilhite, 1999). Suitable drought indices are required to determine the severity and duration of droughts. These drought indices are typically based on rainfall anomalies although a variety of alternatives has been developed (Vicente-Serrano et al., 2010).
Due to the complexity of multivariate analysis, most studies are limited to univariate frequency analysis of severity and duration. To extend this to a bi-variate analysis, a copula can be used to link the univariate marginal distribution functions together (Sklar, 1959). The introduction of copulas for bivariate severity-duration frequency (SDF) analysis has led to new challenges and applications (Shiau et al., 2012). SDF curves aid decision making by providing Recurrence Intervals (RIs) of droughts (Todd et al., 2013), comparing the relationship of droughts with climate variables such as El-Niño, analysing global climate change and mitigation options, analysing regional droughts (Shiau et al., 2012; Shiau and Modarres, 2009), and specifying design droughts for rehabilitating ecosystems (Halwatura et al., 2015a).

However, there is uncertainty involved in the calculation of SDF curves, which can impact the final outcome. Uncertainties potentially stem from the limited length of the observed data record (Halwatura et al., 2015a), the choice of drought index, the level of aggregation of data with respect to the timescale of drought (e.g. 3-month, 6-month or 12-month timescales), and the probability distribution function and copula used to describe the marginal and joint distributions of severity and duration. For reliable SDF analyses, ideally the data needed to calculate an index would be available for a continuous period of at least 30 years (McKee et al., 1993) because the estimation accuracy is sensitive to the observed sample (Hu et al., 2015).

The identification of a suitable drought index is not straightforward as there are currently around 100 drought indices used to quantify different types of droughts (Zargar et al., 2011). Some drought indices mainly focus on particular types of drought, while other indices can be configured to correspond to varying impacts of droughts (Zargar et al., 2011). Similarly, different practitioners use drought indices designed for various temporal scales of drought and thus it is important to understand the timescale relevant for a particular study (Passioura, 2007). Furthermore the distribution functions (gamma, logistic, extreme value, lognormal) and the type of copula (Gumbel, Frank, Clayton) (Shiau, 2006) may need to be chosen based on the region and data set.

The objective of this paper is to examine the uncertainties associated with estimation of SDF curves. We assess the impact of uncertainty on estimates of drought RI by analysing three drought indices (RDI, SPEI, SPI), three time scales (three, six and twelve months), four marginal probability distribution functions (extreme value, logistic, gamma, lognormal), and two types of copulas (Gumbel and Frank). The paper will thus contribute to understanding uncertainty in the calculation of SDF curves and illustrate the need to consider the key sources of uncertainty as part of drought planning and decision making.

2 METHODS

This study builds on a previous study (Halwatura et al., 2015b) of SDF analyses across 11 locations (Figure 1), each representing specific soil and climate combinations across Eastern Australia. Historical observations of monthly total rainfall and potential evaporation (from 1960-2013 ranging from 30–60 years) from weather stations at same sites were used (Bureau of Meteorology, 2013).

2.1 Baseline case

The baseline case used the Reconnaissance Drought Index (RDI3). For a month k, this is defined as:

\[ RDI_{3,k} = \frac{y_{3,k} - \bar{y}_3}{\hat{\sigma}_3}, \]  

(1a)

with

\[ y_{3,k} = \ln \left( \frac{\sum_{j=k}^{k+2} p_j}{\sum_{j=k}^{k+2} PET_j} \right) \]  

(1b)
Where $y_{3,k}$ is a running mean index of wetness beginning at month $k$, $\bar{y}_3$ is the arithmetic mean of $y_3$ over all $k$, $\sigma_3$ is the standard deviation of $y_3$ over all $k$, and $P_j$ and $PET_j$ are rainfall and potential evapotranspiration for month $j$ within each 3-month window (Tsakiris and Vangelis, 2005).

A drought event starts whenever the index becomes negative and the event stops when index becomes positive again, so that the duration ($D$) is the number of consecutive values of index below zero. The severity ($S$) of a drought starting in month $i$ is defined as:

$$S = \sum_{i=1}^{D} |RD|.$$  

(2)

For the baseline case, the gamma and logistic distributions were used to describe the marginal distributions of the annual maxima $S$ and $D$ respectively. This followed satisfactory fits using regression of the reduced variates against the corresponding plotting position estimates. The dependency between $S$ and $D$ was represented by the Frank copula.

The parameters of the two distributions were estimated using all drought events in the respective year and using the regression method the copula was used to join the univariate drought severity and duration into a bivariate joint distribution (Halwatura et al., 2015b). Random drought events were generated from the joint distribution (Halwatura et al., 2015b) (Figure 2). Baseline SDF curves were generated for all 11 sites.

<table>
<thead>
<tr>
<th>Test case</th>
<th>DI</th>
<th>Time scale (months)</th>
<th>Distribution function</th>
<th>Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>RDI</td>
<td>3</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>B</td>
<td>RDI</td>
<td>3</td>
<td>Extreme value (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>C</td>
<td>RDI</td>
<td>3</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>D</td>
<td>RDI</td>
<td>3</td>
<td>Gamma (S) Logistic (D)</td>
<td>Gumbel</td>
</tr>
<tr>
<td>E</td>
<td>RDI</td>
<td>6</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>F</td>
<td>RDI</td>
<td>12</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>G</td>
<td>SPI</td>
<td>3</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
<tr>
<td>H</td>
<td>SPEI</td>
<td>3</td>
<td>Gamma (S) Logistic (D)</td>
<td>Frank</td>
</tr>
</tbody>
</table>

*baseline, DI: drought index

B and C) The best fitting alternatives to the gamma and logistic distributions were the extreme value and lognormal for severity and duration respectively, so these were used instead of the gamma and logistic separately. However Bourke, Cairns, Kingaroy, Quilpie, Wagga (see Figure 1 for locations) had poor correlations $(R^2<0.3)$ with extreme value and lognormal distributions, therefore we did not run the analysis for case B and C for these sites.

D) SDF curves derived using Gumbel copula with a similar fit to Frank copula used for baseline.

Figure 2. Flow chart of the possible uncertainties in the calculation process

2.2 Test cases

After estimating the SDF curves using the baseline case (case A), seven adaptations (cases B-H) to the method were used for comparison at each of the 11 sites. Everything was kept the same as in the baseline case except as noted below (Table 1):
E) Instead of aggregating the data over 3 months as in Equation 1b, a 6 month period was used.
F) A 12-month period was used.
G) Standardized Precipitation Index (SPI) was used instead of RDI. Another widely accepted drought index.
H) The Standardized Precipitation Evapotranspiration Index (SPEI) was used instead of RDI. Another widely accepted drought index.

The SDF curves represent the recurrence interval (T) of drought events exceeding any severity or duration of interest (equation 3) (Halwatura et al., 2015b),

\[ SDF_{x,y} = \frac{1}{p(D>x \text{ OR } S>y)} \] (3)

where, \( D \) is maximum duration of the year and \( S \) is the maximum severity.

Three selected types of drought events (mild: \( S<5 \) or \( D<5 \) months; moderate: \( 5<S<10 \) or \( 5 \text{ months}<D<10 \text{ months} \); extreme: \( 10<S<15 \) or \( 10 \text{ months}<D<15 \text{ months} \)) were selected for further calculations (Figure 3). The average RIs of each case for all sites for the three selected drought types were calculated in table 2. Finally the ratio between the RIs of baseline and selected cases were calculated and for the sake of simplicity, we present only the values of extreme droughts (Table 3).

Figure 3. Three selected ranges of droughts

3 RESULTS AND DISCUSSION

SDF curves are a common drought analysis tool which has been used in many parts of the world for different drought related circumstances. Yet, there are issues related to the reliability of SDF curves due to uncertainties introduced in each step of the calculation. Our results, discussed below, showed that uncertainties in RIs can arise due to choice of time-scale, while other factors such as choice of copula had little impact on RIs; and generally, the variation in RIs between baseline and other cases increased with the extremeness of the drought type (Table 2). The results are discussed further below.

3.1 Distribution function (case B and C)

Bivariate distributions were originally and commonly used for describing correlated hydrologic variables and eventually used for bivariate drought analysis to derive the distribution of drought duration and severity (Shiau and Shen, 2001). In general, studies related to SDF modelling have used the best fitted distribution function for their data set without any explanation why they were best fitted. For the data set used in this study, only the gamma and extreme value (severity), as well as the logistic and lognormal (duration) distributions fitted well (\( R^2>0.5 \)) with observations of severity and duration, respectively. On the other hand, the extreme value distribution (case B) and lognormal distribution (case C) were poorly correlated (\( R^2<0.3 \)) for four sites (Bourke, Cairns, Kingaroy, Quilpie, Wagga) where rainfall is highly seasonal. Our results reveal that the effect on the RIs of changing the distribution functions (cases B & C) was one year for mild droughts and much higher for moderate and severe events (Table 2, Figure 4).

Similar to our results, other studies have found, for severity, the gamma distribution was predominantly identified as the best-fitting distribution regardless of the climatic region (Reddy and Ganguli, 2012; Shiau et al., 2007). For duration, a range of distributions have been used successfully in a range of different climatic regions. For example, Mishra and Singh (Mishra and Singh, 2009); Shiau et al.
(2007) used extreme value, logistic, and lognormal distributions for tropical and arid climates. While, Vicente-Serrano (2006) used lognormal and extreme value distributions in temperate regions. Finally, the extreme value distribution has been used in many climatic regions (Ganguli and Reddy, 2012; Lee et al., 2013; Reddy and Ganguli, 2012). None of these distribution functions showed better correlations for particular climates (Shiau et al., 2007).

### Table 2. Average RIs of each case for all sites for the three selected ranges of drought (case A denotes the baseline).

<table>
<thead>
<tr>
<th>Drought type</th>
<th>Selected cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A  B  C  D  E  F  G  H</td>
</tr>
<tr>
<td>Mild</td>
<td>3  2  2  3  2  3  3</td>
</tr>
<tr>
<td>Moderate</td>
<td>15 22 21 15  8  2  9 13</td>
</tr>
<tr>
<td>Extreme</td>
<td>60 196 189 56 16  3 24 41</td>
</tr>
</tbody>
</table>

### 3.2 Copula (case D)

The use of copulas provides an alternative to using more traditional joint distribution functions such as the bivariate normal or log-normal distribution functions, and often provide a better fit to the data (Shiau and Modarres, 2009; Sklar, 1959). Of the copulas available for bivariate analysis Gumbel and Frank copula are often considered to be the best for modelling the joint dependence structure of drought variables (Ganguli and Reddy, 2012; Lee et al., 2013; Reddy and Ganguli, 2012; Shiau, 2006). Our results showed that for each of the three selected drought types (mild, moderate, extreme) there was no difference between the RIs derived from the Frank copula (baseline case A) and those derived from the Gumbel copula (case D) (Table 2); and no particular copula was better for a specific climate. Other studies selected copulas independently of their respective climate, e.g. the Clayton copula has been widely used for temperate climates, while the Galambos copula and the Empirical copula have been used for arid reigns. All these copulas have been used in tropical climates (Liu et al., 2011; Mirabbasi et al., 2012; Shiau et al., 2012; Shiau et al., 2007).

### Table 3: The ratio of the RIs between the selected cases and the baseline case for the extreme droughts.

<table>
<thead>
<tr>
<th>Location</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brigalow</td>
<td>0.3</td>
<td>0.3</td>
<td>1.0</td>
<td>2.9</td>
<td>15.7</td>
<td>1.4</td>
<td>0.9</td>
</tr>
<tr>
<td>Bourke</td>
<td>*</td>
<td>*</td>
<td>1.0</td>
<td>3.3</td>
<td>6.7</td>
<td>1.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Brisbane</td>
<td>0.4</td>
<td>0.4</td>
<td>1.0</td>
<td>5.5</td>
<td>20.0</td>
<td>1.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Cairns</td>
<td>*</td>
<td>*</td>
<td>1.0</td>
<td>0.2</td>
<td>40.3</td>
<td>2.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Mount Isa</td>
<td>0.1</td>
<td>0.1</td>
<td>1.0</td>
<td>1.8</td>
<td>16.5</td>
<td>1.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Kingaroy</td>
<td>*</td>
<td>1.1</td>
<td>1.0</td>
<td>6.8</td>
<td>16.4</td>
<td>2.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Melbourne</td>
<td>0.1</td>
<td>0.1</td>
<td>1.0</td>
<td>2.7</td>
<td>12.7</td>
<td>2.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Quilpie</td>
<td>*</td>
<td>*</td>
<td>1.0</td>
<td>4.7</td>
<td>20.3</td>
<td>6.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Sydney</td>
<td>0.2</td>
<td>0.2</td>
<td>1.0</td>
<td>3.7</td>
<td>28.0</td>
<td>2.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Wagga</td>
<td>*</td>
<td>*</td>
<td>1.0</td>
<td>1.5</td>
<td>10.7</td>
<td>2.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Weipa</td>
<td>0.3</td>
<td>0.3</td>
<td>1.0</td>
<td>2.9</td>
<td>15.7</td>
<td>1.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*Sites showed poor correlations (R^2<0.3) with extreme value and lognormal distributions, therefore we did not run the analysis for case B and C for these sites.*
3.3 Timescale (cases E and F).

The time scale of a drought refers to the time lag between the start of a drought and its impact. An index may be calculated with time scales which vary from months (short term drought) to years (long term droughts). The time-scale should be selected based on the type and the purpose of the study. For example, the time scale of a drought of interest to meteorologists may vary from months to years, while plant physiologists may be interested in the number of days of water deficit (Passioura, 2007). We found that when using the 6 month time-scale, at Kingaroy the RI was 7 times the value found for the baseline case, while the RIs of Quilpie and Brisbane were around 5 times the baseline value. The twelve month time scale (case F) had higher ratios than the 3 month timescale with the RIs of Cairns and Sydney values 40 and 28 times higher than the baseline respectively. For moderate and extreme droughts, RIs of longer time scales (case E & F) showed much lower values than the baseline. The results also reveal that the difference between the SDF curves of the same drought index increases with time scale (Table 2). For long time scales (12 months; case F) the difference between RDI3 and RDI12 increases with increasing drought severity and duration (extreme droughts). The difference in RIs between RDI3 and RDI12 was more than 10 times for extreme droughts (Table 3). Therefore our results showed that the variation of the RI is primarily dependent on the selected time scale. However, droughts may be estimated more reliably by some indices at specific time scales, for example, the SPEI detects annual drought events more reliably than other time scales, whereas RDI has been found more suitable for detecting droughts on 3 - 6 month time scales (Banimahd and Khalili, 2013; Halwatura et al., 2015b).

3.4 Drought index (cases G and H)

The selection of a drought index that characterizes drought levels by integrating one or several meteorological or hydrological data is difficult as there are more than 100 drought indices. These indices are more practical and informative than raw data such as rainfall, evaporation, river flow (Mishra and Singh, 2010). However two different drought indices using the same data may deliver different results. Our data showed that the baseline calculated with RDI was more similar to SPEI (case H – calculated using rainfall and potential evaporation) than SPI (case G - calculated using rainfall) (Table 3). This was the case for most locations except for SPI values in Quilpie (arid location), which showed 7 times higher RIs than the baseline. These patterns in differences between drought indices have also been observed by Halwatura et al. (2015b). Several studies have showed SPI, SPEI and RDI RIs are similar for different time scales and climates. For example, Halwatura et al. (2015b); (Khalili et al., 2011) found in tropical climates SPI3 and RDI3 were correlated and in our study the differences between RDI, SPEI and SPI for tropical Cairns and Weipa were small. Spinoni et al. (2013) stated that in temperate climates SPI, SPEI, and RDI are highly comparable. However, for mild droughts SPEI (case H) had the same RIs as the baseline, except for temperate Sydney, Kingaroy and Brisbane (Peel et al., 2007), with RIs twice as high as the baseline. For SPI (case G) Sydney, Melbourne, Kingaroy and Wagga Wagga which have the similar climatic conditions (temperate-without dry season) the values were three times higher than the baseline. Therefore, uncertainty associated with the choice of drought indices depends on the climatic conditions of the location. However the selection will also be restricted by the availability of data as some indices require other climatic records such as evaporation (Tsakiris and Vangelis, 2005).

Figure 4. Recurrence intervals T* (years) of drought events of any severity or duration of interest based on selected cases for Melbourne
4 CONCLUSION

Understanding the recurrence intervals of droughts and their severity and duration is a key aspect of drought management. However, the uncertainty associated with input data and the parameterization of SDF curves will affect the utility of these methods for estimating drought RIs. The results show that the longer timescales tend to have the greatest influence on calculated RIs while the chosen copula was the least influential parameter, with an almost negligible effect. Therefore comprehending the uncertainty involved in developing SDF curves may aid the applicability of design droughts in rehabilitation of drought affected ecosystems.

5 REFERENCE


