We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Hydrological Drought Analysis Based on Copula Theory

Jaewon Kwak, Soojun Kim, Duckhwan Kim and Hungsoo Kim

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/64244

Abstract

Drought has been a more frequent phenomenon of major concern all over the world. From the perspective of water resources management, one of the biggest problems associated with drought analyses is a lack of quantitative estimation for the target drought amount. The objective of this study is to examine the establishing process for the severity-duration-frequency (hereafter referred as “SDF”) curves on climate change. The standardized truncation level that defines hydrological drought was estimated and a bivariate frequency analysis for drought duration and severity was derived. The SDF curves were also estimated. The methodology suggested in this study could be used as elementary data for water resources managements.

Keywords: hydrological drought, frequency analysis, run theory, copula

1. Introduction

In recent times, drought has become one of the major concerns because it has occurred more frequently all over the world. Droughts are caused by the shortage of rainfall, affecting water resources in both urban and rural areas, and they cause the greatest damage among all the natural disasters [1]. For the Korean peninsula, a drought occurs approximately once every 2 years and the costs and losses due to drought are also increasing dramatically [2]. Generally, drought has been grouped by type as meteorological, hydrological, agricultural, and socio-economic drought [3]. Among these, hydrological drought causes real damage to economic sectors, because it is an actual deficiency of water in hydrological storage systems, such as streamflow, reservoir and lake levels, groundwater, power generation, irrigation, and recreation. In order to develop measures for the mitigation of hydrological droughts, it is important to quantify...
the probabilistic characteristics of droughts. Therefore, there are many ongoing investigations for hydrological droughts quantitative estimation and considering future climate change.

A number of studies have been conducted on the analysis and estimation of hydrological drought. These studies can be divided into several topics. Some studies have employed hydrological drought indices for quantitative estimation, such as cumulative streamflow anomaly that a cumulative departure of streamflow from mean conditions [4–6], surface water supply index (SWSI) that is a suitable measure of hydrological drought for the mountainous region, where snow contributes significantly to the annual streamflow [4, 7, 8], or Palmer hydrological drought severity index (PHDI) that uses the identical water balance with soil moisture model [5, 7, 9, 10]. Other studies also conducted hydrological drought analysis using deterministic approach such as the extreme dependency score that is an informative assessment of skill in deterministic forecasts [11], or stochastic model [12, 13]. The other, other studies were conducted on probabilistic approach for hydrological drought monitoring [14–16], or forecasting [17–19]. In detail, since Yevjevich [20], many studies on univariate analysis have been conducted based on the run theory [21]. These studies assumed run length and sum to be independent identically distributed random variable and have analytically derived statistical characteristic of hydrological drought, such as probability distributions, return periods, accumulated deficit, and moments of drought durations [22]. By contrast, the studies about multivariate drought analysis are relatively small, because they require considerably more data and sophisticated process, which are limited in their applicability [23]. Some studies have employed bivariate gamma, exponential, and extreme distribution for drought analysis, but they have limitations that drought characteristics have to have same distribution type. Therefore, many studies have been undertaken, because they can be overcome by the copula theory, which was suggested by Sklar [24]. The drought surmised that the increased with climate change, there are also many on-going studies about drought are conducted [26] with Global Climate Models and Regional Climate Model (hereafter referred as ‘GCM’ and ‘RCM’) [25–31]. Considering that the GCM/RCM is one of the effective ways for future projection, drought frequency analysis based on the copula theory and GCM/RCM could be a good alternative for water resources managements or planning.

The objective of this chapter is therefore to briefly examine the establishing process for the SDF curves on climate change based on the studies of Kwak et al. [32, 33]. For this, the joint probability distribution of drought characteristics using copula theory was examined, and established that the derive process the drought SDF curves for presents and future projected period using GCM/RCM. These established concepts and processes could be the basic methodology for water resources planning.

2. Methodologies

2.1. Drought definition with run theory

Generally, drought is a prolonged deficiency of precipitation including snow, a deficiency that results in water shortage for some activity or for some group, or a period of abnormally dry
weather sufficient to cause a serious of hydrological imbalance [34]. Drought has been defined in a number of ways and hydrologic drought is related to below-normal streamflow, lake, and groundwater levels. Yevjevich [20] defined drought using the run theory and suggested a way to calculate drought variables (drought duration, severity, and inter-arrival time).

Figure 1. Drought characteristics using the run theory.

As shown in Figure 1, each drought variables are defined by truncation level that can be determined as a constant value or a time-varying function. Specifically, according to the truncation level $x_0$, the time period when the streamflow falls below $x_0$ defines “duration (D1, D2, ...)” and the accumulated shortage amount of the streamflow during drought duration and the interval time period between two drought occurrences defines drought severity (S1, S2, ...) and inter-arrival time (I1, I2, ...), respectively. Therefore, the proper truncation level is the most challenging problem, and it has been widely used to estimate statistical drought model and analysis [25, 26, 35–37] due to their applicability.

2.2. Copula theory

One of the challenges of the multiple variable analysis of drought is the different type of distribution due to their characteristic that variables are highly correlated with each other [32]. Therefore, the copula function that was suggested by Sklar in 1959 [24] was thought to be an effective alternative to consider the dependence structure between drought variables [22, 23]. The copula function was used to measure the correlation between multiple variable such as the coefficient of correlation, but it also has the advantage that can be used in any type of distributions. Especially, it is more appropriate than the correlation coefficient in the case of variables, which tend to indicate the same directionality [33].

As Sklar’s theorem, there is a copula function $C$ with a dependence structure of probability distribution $F(x_1, ..., x_n)$ with $n$-dimensional function having marginal distributions $F_1(x_1), ..., F_n(x_n)$, and it can be represented as
where \( F_i(x) \) is the marginal distribution of each variable. When each marginal distribution \( F(x_1, x_2, \ldots, x_n) \) is satisfied to be continuous random, the copula function also can be represented as

\[
C(x_1, x_2, \ldots, x_n) = F(F_1^{-1}(x_1), F_2^{-1}(x_2), \ldots, F_n^{-1}(x_n))
\]

(2)

When \( u_i = F_i(x) \in [0,1] \) and \( F_i^{-1} \) is the inverse function of \( F_i \), the partial differentiation of both Eqs. (1) and (2) with respect to \( x_1, \ldots, x_n \) can be yielded as follows:

\[
f(x_1, \ldots, x_n) = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \prod_{i=1}^{n} f_i(x_i)
\]

(3)

\[
c(u_1, u_2, \ldots, u_n) = \frac{f(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n))}{\prod_{i=1}^{n} f_i(x_i)}
\]

Using Eq. (3), the multivariate probability density function that considers the relationship between marginal probability distribution can be obtained. It also shows that the copula function is the ratio of joint probability distribution function and the product of marginal probability distributions (see Table 1).

<table>
<thead>
<tr>
<th>Copula Family</th>
<th>Copula function, ( C(F_1(x), F_2(x)) )</th>
<th>Generator function ( \psi(t) )</th>
<th>Generator function ( \psi(a) )</th>
<th>Parameter (( \alpha ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>( \max\left[F_1(x_1)^{-1} + F_2(x_2)^{-1} - 1; 0\right] )</td>
<td>( \frac{t}{\alpha t - 1} )</td>
<td>( \frac{\exp(-at) - 1}{\exp(-1) - 1} )</td>
<td>( \alpha \in [-1, \infty] )</td>
</tr>
<tr>
<td>Frank</td>
<td>( -\frac{1}{\alpha} \log\left(1 + \frac{\exp(-aF_1(x_1)) - 1}{\exp(-1) - 1} \right) )</td>
<td>( \exp(-at) )</td>
<td>( \exp(-at) )</td>
<td>( \alpha \in \mathbb{R} )</td>
</tr>
<tr>
<td>Gumbel</td>
<td>( \exp\left(-((\log(F_1(x_1))^\alpha + (\log(F_2(x_1))^\alpha)\right)^{1/\alpha} )</td>
<td>( -\log(t)^\alpha )</td>
<td>( \log(t)^\alpha )</td>
<td>( \alpha \in [1, \infty] )</td>
</tr>
<tr>
<td>Independence</td>
<td>( F_1(x_1)F_2(x_2) )</td>
<td>( -\log(t) )</td>
<td>( -\log(t) )</td>
<td>( \alpha \in [1, \infty] )</td>
</tr>
<tr>
<td>Ali Mikhail Haq</td>
<td>( \frac{F_1(x_1)F_2(x_2)}{1 - a(1 - t)(1 - F_2(x_2))} )</td>
<td>( \log\left[\frac{1 - a(1 - t)}{t} \right] )</td>
<td>( \log\left[\frac{1 - a(1 - t)}{t} \right] )</td>
<td>( \alpha \in [-1, 1] )</td>
</tr>
</tbody>
</table>

Table 1. Archimedean copula family; \( C \) is the copula function, \( t \) denotes the drought event, \( a \) is the copula parameter, and \( F_1 \) and \( F_2 \) denote marginal distributions [38].
The occurrence probability with copula function can be represented to be the return period, that is, the average time of occurrence of hydrological events of certain intensity. The return period was calculated from univariate frequency analysis using the following equations:

\[
1/T = \frac{E(L)}{1 - F_1 - F_2 - \ldots - F_n + C(F_1, F_2, \ldots, F_n)}
\]

where \( L \) is the inter-arrival time between drought events, \( E(L) \) is the average inter-arrival of occurrences, and \( F_1 \) and \( F_2 \) are the cumulative probability distribution functions of drought duration and severity, respectively.

### 3. Drought frequency analysis

#### 3.1. Study area and data

The upstream of the Namhan River is a main stream of Han River, with a drainage basin of 2,442.22 km\(^2\), located in the middle-eastern part of Korea. The upstream of the Namhan River is unregulated with flows between 8.0 and 1,000.0 m\(^3\)/s at the Yeongwol water-level station, which is the outlet of the study area, and shows extreme flow rate in the typhoon season during June to September. The streamflow data were obtained from 1967 to 2008 with flow meter sampling at hourly interval. Also, other hydrological properties of the study area were extracted from a 3 × 3-km grid digital elevation model (DEM), land-cover, and land-use map [39]. The collected streamflow data are shown in **Figure 2**. Drought duration, severity, and inter-arrival time were defined for streamflow data from the upstream of the Namhan River using the run theory.

![Figure 2. Monthly streamflow upstream of Namhan River, Korea (1967–2008).](image)
3.2. Derive drought event based on run theory

The truncation level is defined as a fixed constant or a time function of streamflow that falls below the relevant level. It is one of the important challenges to estimate the proper bivariate frequency analysis because it determines the drought characteristics. Some studies employed the monthly mean value of streamflow, which is widely used due to their applicability, as the truncation level [23, 40–43]. The median value that has advantage with abnormal extreme value is also widely adopted [44, 45]. However, the mean and median values of streamflow are not proper, because the ratio between the minimum and maximum flow is over 300 in the river of Korea. In this case, the standardized truncation level should be a good alternative [46]. Generally, 0.5 (median value), 0.7, 0.8, 0.9, and 0.95 are commonly employed as the standardized truncation level, and Korea is adopting the concepts of Low Flow and Minimum Flow in the field of water resources planning; it is equal as 0.75 and 0.97 [47]. Also, the monthly streamflow in Korea shows significantly seasonal effect, so there are less or more occurrences of droughts than expected. Therefore, the truncation levels are defined for each month with monthly standardized truncation level (see Figures 3–5).

Each standardized truncation level was examined in comparison with major drought records in South Korea [48], and 0.75 is selected as the proper truncation level. The mean monthly truncation level was considered relatively improper, because it led to too large and deeper drought events. The defined droughts are 2.12 months of average drought duration, 22-mm/month severity, and 8.8 month of inter-arrival time (see the Figure 3). The events from December 1981 to July 1982 were estimated as the most severe drought with a 109.5-mm severity and an 8-month duration, and their damage were estimated about 643 million USD as of the year 2000 [48].

![Figure 3](https://example.com/figure3.png)

Figure 3. Drought event of the upstream of the Namhan River (1967–2008) with 0.75 standardized truncation level for each month.
3.3. Bivariate drought analysis based on copula

Each marginal distribution of drought duration and severity is essential to estimate bivariate frequency analysis. The drought duration was determined to have the “exponential” distribution, and the drought severity the “gamma” distribution with 95% confidential level with the Probability Plot Correlation Coefficient (PPCC) [49] distribution goodness-of-fit test. Furthermore, the parameters of the Archimedean family copulas (Frank, Clayton, and Gumbel) were estimated by the method of moments according to their relationship with Kendall’s tau [50], which was found adequate for estimating parameters for small sample sizes [51]. Also, the minimum quadratic distance ($L^2$) between empirical and theoretical values of
the K criterion, which described the most appropriate copula [51], was estimated for each copula (see the Figure 4).

The Gumbel copula, which generally fitted well throughout \( (L^2 = 0.0138) \), was selected for bivariate drought analysis for the study area. The copula parameter was estimated as 3.599.

With the bivariate probabilities, it is possible to combine the drought duration and severity, and express them in terms of the same return period. For example, the severest drought lasted for 8 months from December 1981 to July 1982 and was 75.51-mm severity, which is approximately 280 years of return period (see Figure 5). The conventional method of deriving intensity-duration-frequency (IDF) curve was applied to drought events and was compared with the copula method, and the results are shown in Figure 6, and it can be used as the elementary data for water resources planning. For instance, if a decision maker or an agency determined a 3-month duration with a 20-year return period, then the design deficiency of a dam or a reservoir is 26.0 mm, which is about \( 63.6 \times 10^6 \text{ m}^3 \).

![Figure 6. Drought SDF curve for the upper Namhan River basin.](image)

4. Bivariate drought analysis on climate change and SDF curve

One of the difficulties to analyze the hydrological drought is streamflow data, which is the result from hydrological response of basins. GCMs representing physical processes in the atmosphere, ocean, cryosphere, and land surface simulate the response of the global climate system [52]. Therefore, they need other techniques to obtain projected streamflow in the future, and long-term hydrological model using meteorological data was commonly used. Also, there
are many climate models in use, and Korea Meteorological Administration (KMA) RCM climate model was selected as the suitable GCM by reviewing the applicability of 25 climate models provided by Intergovernmental Panel on Climate Change (IPCC) [53]. Also, socioeconomic scenarios are also needed to simulate meteorological data, and A1B climate change scenario was selected as the suitable scenario [54]. The meteorological data until 2100 years on the upper Namhan River basin were projected using climate model and scenarios. The modified TANK model [55] was constructed as the long-term hydrological model, and was calibrated and validated with observed meteorological data. Based on the projected meteorological data and hydrological model, the streamflow data until 2100 years were simulated (see Figure 7).

To analyze short- and long-term trends of the drought, each case was defined as every 30 years; Obs Case for 1967–2007; Case 1 for 2011–2039; Case 2 for 2040–2069; Case 3 for 2070–2100. The
projected monthly mean streamflow is shown in Figure 8, and it presents that the streamflow is highly increased from March to June and decreased from September to December.

The projected drought event can be obtained through the truncation process with simulated streamflow and 0.75 standardized truncation levels. Moreover, it can also be estimated as the return period and SDF curves with the same process as described in Section 3.

The SDF curves that are derived from climate change model can be used for water resources planning as the elementary data. For example, in Figure 9, when the agency determined for design drought has a 30-year return period and a 3-month duration, the design amount is 35.80 mm/month. Therefore, the target storage volume for dams, reservoirs, and other hydrological countermeasures is 87.6 million m³.

![Figure 9. SDF curves for each case.](image)

Brief steps to establish SDF curve on climate change are shown as follows:

**Step 1.** Collecting a hydrological, hydro-climatic, and drought data for target basin.

**Step 2.** Finding proper standardized truncation level for basin with observed data.
Step 3. Finding a proper distribution type with goodness-of-fit test.

Step 4. Selecting proper climate model.

Step 5. Constructing long-term hydrological model and conducting calibration and validation.


Step 7. Extracting drought event based on the standardized truncation level.

Step 8. Bivariate frequency analysis with copula method.


5. Conclusions

This study briefly examines the establishing process for the SDF curves on climate change. The standardized truncation level was estimated with actual drought damage, and bivariate frequency analysis for drought duration and severity was estimated, and the SDF curves were also derived. Also, the SDF curves until 2100 years were derived based on the climate change models and scenarios, and long-term hydrological models are constructed and validated. The suggested method in this study can be used for water resources managements and planning as the quantitative indicator.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) and by a grant funded by the Korean government (MEST; No. 2011-0028564).

Author details

Jaewon Kwak¹, Soojun Kim²*, Duckhwan Kim³ and Hungsoo Kim³

*Address all correspondence to: soojun78@gmail.com

¹ Forecast and Control Division, Nakdong River Flood Control Office, Busan, Korea

² Columbia Water Center, Columbia University, New York, NY, USA

³ Department of Civil Engineering, Inha University, Incheon, Korea
References


