Characterizing extreme rainfalls and constructing confidence intervals for IDF curves using Scaling-GEV distribution model

Myeong-Ho Yeo | Van-Thanh-Van Nguyen | Theodore A. Kpodonu

1Water and Environmental Research Institute of the Western Pacific, University of Guam, Guam
2Department of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada
3Auckland Council, Health Waters, Infrastructure & Environmental Services, Auckland, New Zealand

Correspondence
Myeong-Ho Yeo, Water and Environmental Research Institute of the Western Pacific, 303 University Drive, University of Guam, Mangilao, GU 96923, USA.
Email: yeom@triton.uog.edu

Abstract
This article compares the performances of three fitting methods (SLmom, S1NCM, and S3NCM) to account for temporal characteristics of Annual Maximum Precipitations (AMPs) on daily and sub-daily time scales using scaling General Extreme Value (GEV) distribution at a local site. Based on simple scaling properties of AMPs, the temporal downscaling model (called Scaling-GEV) with parameter estimation methods are used to estimate sub-daily AMPs from observed daily data. The feasibility and accuracy of the suggested method were assessed using rainfall data available from Dorval in Quebec (Canada) and Seoul (South Korea) for the period 1961–1990. Presence of simple scaling properties of AMPs for two stations has shown that it is feasible to use the temporal downscaling method for describing the linkage between AMPs of different time scales. Numerical and graphical analyses revealed that the Scaling-GEV distribution by the Three-Non central moments (NCM) method (S3NCM) provides the most accurate estimates compared to observed data amongst three fitting methods. In addition, this study suggested a modified bootstrap technique to determine confidence intervals (CIs) CIs of extreme rainfall series using the simple scaling properties of extreme rainfalls and only daily AMPs. Although the CIs were constructed by only daily AMPs and the simple scaling properties, the observed sub-daily AMPs are generally within the 95% CI estimated.

KEYWORDS
extreme rainfalls, intensity-duration-frequency relations, generalized extreme value distribution, scale-invariance

1 INTRODUCTION

The estimation of the extreme rainfall for a given duration and for a selected return period is often necessary for the planning and design of various hydraulic structures such as dams, reservoir, storm sewers, and culverts. For a site for which sufficient rainfall data are available, a frequency analysis of annual maximum rainfalls can be performed. Results of this analysis are often summarized by “intensity-duration-frequency” (IDF) relationships for a
Several probability models have been developed to describe the distribution of extreme rainfalls at a single site (Buishand, 1989; Wilks, 1993; Zalina et al., 2002). Unfortunately, these models are accurate only for the specific time frame associated with the data used. In other words, the inferences using traditional methods are applicable only to the particular time scale of the data used and the site it was predicted for. This normally results in the diminishment of the models’ predictive capability for other time durations. Conventional models (Hosking, 1990) cannot capture extreme rainfall on sub-daily time scales that are not measured. This has therefore necessitated the formulation of models that could statistically and simultaneously match various properties of the extreme rainfall process at different levels of aggregations. To statistically describe the relationship of annual maximum precipitations (AMPs) for certain duration to those for different duration, scale invariant concept, a type of fractal mathematics, has been proposed and applied to different regions (Bara et al., 2009; Burlando and Rosso, 1996; Menabde et al., 1999; Nguyen, 2004; Nguyen et al., 2002; Yeo, 2013).

The main practical implication of the models based on scaling invariance properties is that we could infer the statistical properties of AMP process on a shorter time scale (e.g., hourly maximum rainfall) using those at the longer time scale (e.g., available daily maximum rainfall). A major advantage of such procedure involves the parsimonious parameter estimation since these models would normally require a much smaller number of parameters, while traditional models need different sets of parameters for each particular time scale of the rainfall series considered (Burlando and Rosso, 1996; Kotegoda and Rosso, 1997). Menabde et al. (1999) developed a simple scaling methodology to build the association between extreme rainfall intensities of two different durations based on the following equation:

\[ X_d = \left( \frac{d}{D} \right)^{-\beta} \cdot X_D \]  

(1)

in which \( d \) denotes equality of probability distribution, \( X_d \) is the rainfall intensity for the shorter duration \( d \), \( D \) is the longer time scale (i.e., 24 h), and \( \beta \) is a scaling exponent. Once the scaling invariance (so called “scaling”) properties of AMPs are built, the statistical properties for shorter duration \((d)\) can be derived using the scaling exponents and the ratio of a shorter time scale to a longer time scale.

Scale-invariance properties of extreme rainfalls have been used for temporally downsampling sub-daily extreme rainfall from daily extreme rainfall in association with several probability distributions. Burlando and Rosso (1996) examined the use of simple and multi-scaling properties of annual maximum precipitation series and proposed the use of two-parameter log-normal distribution for constructing rainfall depth-duration-frequency curves, while others used the Gumbel distribution for deriving IDF curves for sub-daily durations from daily extreme rainfall records (Choi et al., 2019; Menabde et al., 1999; Vu et al., 2018). Also, Generalized Extreme Value (GEV) was used to describe the temporal relationships between extreme rainfall intensities for different durations in order to describe quantiles lying on distribution tail (Bougadis and Adamowski, 2006; Nguyen et al., 2008; Blanchet et al., 2016). In these studies, parameter estimation methods for scaled-GEV distributions were different from each other. This study illustrates mathematically possible estimation methods for the scale-invariant GEV (called Scaling-GEV) models and evaluate their performances.

Recently, climate variability and change have been recognized to have important impacts on the hydrologic cycle at different temporal and spatial scales. The temporal scales could vary from a very short time interval of 5 min (for urban water cycle) to a yearly time scale (for annual water balance computation). The spatial resolutions could be from a few square kilometres (for urban watersheds) to several thousand square kilometres (for large river basins). General Circulation Models (GCMs) have been commonly used to assess these impacts since these models could describe reasonably well the main features of the distribution of basic climate parameters at global scale. However, outputs from these models are usually at resolutions that is too spatially (generally >200 km) and temporally (daily basis) coarse thus are not suitable for many hydrological impact studies at the regional or local scale. The combination of spatial downscaling models and the scale invariance properties of annual maximum precipitation has been applied for climate change impact studies (Bougadis and Adamowski, 2006; Nguyen et al., 2008; Vu et al., 2018; Yeo, 2013).

Gumbel distribution (or Extreme Value Type 1; EV1) has been used to construct IDF curves for sub-daily durations with scaling or scale-invariance properties in many studies (Aronica and Freni, 2005; Nhat et al., 2008; Vu et al., 2018; Yu et al., 2004). It is because the parameter estimation procedure of distributions for sub-daily or sub-hourly time scales is relatively easy. However, the
application of Gumbel distribution to hydrological extreme values may lead the risk of underestimation as a result of the reduced flexibility in the tail area because the shape parameter of Gumbel distribution is assumed as constant with zero (Koutsoyiannis, 2004). Therefore, we have applied the GEV or Fréchet (EV2), which has a positive shape parameter value, distribution and scale invariance properties to estimate sub-daily AMPs of rain gauges from Canada and South Korea (Nguyen et al., 2002, 2008; Yeo, 2013). In addition to the distribution, this study performed the comparison study between three different estimation methods for sub-daily GEV distribution using scaling properties.

Furthermore, in hydrological frequency analysis, selected distribution with limited sample size is used to estimate quantiles corresponding to several return periods greater than the length of data (Kite, 1974; Nguyen et al., 2002). The uncertainty of estimated quantiles from the distribution was the result of insufficient data size, the procedure for selecting appropriate probability distribution, and the estimation of parameters of the selected distribution. Confidence intervals (CIs) have been then used to indicate the uncertainty of quantiles. Approaches to calculate CIs can be generally classified into two broad categories: parametric and non-parametric (Burn, 2003; Kite, 1974). Parametric approach is based on the method of moments (MOM) and the assumption of normal distribution of extreme quantiles. Due to the assumption, the accuracy of CIs is dependent on sample size and how well the data follow normal assumption. However, non-parametric approaches calculate CIs using resampling techniques such as bootstrap and jackknife methods. Because there is no requirement for any assumption and distribution, it is easy to apply this for determining CIs of extreme quantiles. However, the application of bootstrapping technique fails when estimating extreme values from the distribution since resampling approach cannot provide bigger quantile than the maximum value and lower quantile than the minimum one Chernick (2011).

As a result of climate change, temporal downscaling models are associated with ensembles of multiple climate change scenarios for projecting future IDF relationships and reflecting the uncertainties resulting from the future climate simulations (Fadhel et al., 2017; Lu et al., 2015; Mailhot et al., 2007; Vu et al., 2018). However, with regard to the uncertainty sources, these methods take into consideration only the uncertainty source from the climate change models, not the sources by other sources such as limited sample size and parameters. Furthermore, downscaling modelling based on the scale-invariance properties can be regarded as “information propagation” because the parameters for shorter durations are estimated by those for the longest duration (e.g., daily). The parameter estimation procedure automatically imparts uncertainty on the value of the estimation from daily distribution to downscaled distributions. In this study, we therefore suggest the modified bootstrap approach for covering the uncertainty coming from the downscaled distributions for short-durations.

The main objectives of this study are to propose parameter estimation methods for GEV distribution for sub-daily AMPs and the alternative methods for estimating the distributions of extreme rainfall and for calculating CIs for AMP series, respectively. The remainder of this article is organized as follows: Section 2 describes the parameter estimation methods and the alternative CI method. Section 3 provides the information about the data used in this study. In Section 4, we present the numerical application to two stations from distinct climatic regions. Finally, our conclusions are presented in Section 5.

2 | MODEL DEVELOPMENT

2.1 | Temporal downscaling model using Scaling-GEV distribution

In this study, scale invariance concept is used to describe statistical relationship between parameters of the selected distribution of extreme rainfall amounts for two different time scales. More specifically, the parameters of distributions for short duration rainfall series can be estimated from those for longer duration rainfall using the proposed scaling properties, and then rainfall intensities corresponding to short duration can be calculated. Since the present approach can downscale extreme intensities from rainfall data available for longer duration, this method is called temporal downscaling method in this study.

2.1.1 | Estimation of parameters of GEV using NCMs

The GEV distribution has been commonly used to describe the distribution of extreme rainfalls for different durations and to construct the IDF curves (NERC, 1975; Schaefer, 1990). The cumulative distribution function, \( F(x) \), for the GEV distribution is given as:

\[
F(x) = \exp \left( -\left(1 - \frac{x - \xi}{\alpha} \right)^{\frac{1}{\kappa}} \right) \quad (\kappa \neq 0)
\]  

(2)

where \( \xi, \alpha, \) and \( \kappa \) are the location, scale, and shape parameter, respectively. The \( k \)-th order of non-central
moments (NCMs), \( \mu_k \), for the random variable \( X \) is given by the following equation:

\[
\mu_k = E[X^k].
\]  

(3)

Pandey (1997) introduced the relationship between the NCMs and parameters of the GEV as the following:

\[
\begin{align*}
\mu_k &= \left( \xi + \frac{\alpha}{\kappa} \right)^k + (-1)^k \left( \frac{\alpha}{\kappa} \right)^k \Gamma(1+k\kappa) \\
&\quad + \kappa \sum_{i=1}^{k-1} (-1)^i \left( \frac{\alpha}{\kappa} \right)^{k-i} \left( \xi + \frac{\alpha}{\kappa} \right)^{k-i} \Gamma(1+i\kappa)
\end{align*}
\]  

(4)

in which \( \Gamma(\cdot) \) is the gamma function. Therefore, it is possible to estimate parameters (\( \xi, \alpha, \) and \( \kappa \)) of GEV distribution by the MOM using first three NCMs as shown as the followings:

\[
\begin{align*}
\mu_1 &= \left( \xi + \frac{\alpha}{\kappa} \right) - \frac{\alpha}{\kappa} \Gamma(1+\kappa) \\
\mu_2 &= \left( \xi + \frac{\alpha}{\kappa} \right)^2 + \frac{\alpha^2}{\kappa^2} \Gamma(1+2\kappa) - 2 \left( \frac{\alpha}{\kappa} \right) \left( \xi + \frac{\alpha}{\kappa} \right) \Gamma(1+\kappa) \\
\mu_3 &= \left( \xi + \frac{\alpha}{\kappa} \right)^3 - \frac{\alpha^3}{\kappa^3} \Gamma(1+3\kappa) \\
&\quad + 3 \left( \frac{\alpha}{\kappa} \right)^2 \left( \xi + \frac{\alpha}{\kappa} \right) \Gamma(1+2\kappa) + \left( \frac{\alpha}{\kappa} \right)^3 \Gamma(1+\kappa).
\end{align*}
\]  

(5)

The quantiles (\( X_p \)) can be calculated by the inverse distribution function as follows:

\[
X_p = \xi + \frac{\alpha}{\kappa} \left\{ 1 - [-\ln(p)]^{\frac{1}{k}} \right\}
\]  

(8)

where \( p \) is the exceedance probability of interest and \( \tau \) is the return period.

2.1.2 Scaling-GEV model

The proposed temporal downscaling method is based on the concept of scale-invariance (or scaling). By definition, a function \( f(t) \) is scaling if \( f(t) \) for the \( t \) duration of precipitation is proportional to the scaled function \( f(\lambda t) \) for the \( \lambda t \) duration with all positive values of the scale factor \( \lambda \) (Feder, 2013). That is, if \( f(t) \) is scaling then there exists a function \( C(\lambda) \) such that:

\[
f(t) = C(\lambda) f(\lambda t)
\]  

(9)

With a proportional relation, the ratio of \( f(t) \) to \( f(\lambda t) \) with respect to the scale factor \( \lambda \) can be readily shown (see Data S1) that:

\[
C(\lambda) = \lambda^{-\beta}
\]  

(10)

in which \( \beta \), called a scaling exponent, is a constant for a local site, and that

\[
f(t) = t^\beta f(1)
\]  

(11)

Hence, the NCMs (\( \mu_k \)) of order \( k \) can be expressed by the function of duration \( t \) as follows:

\[
\mu_k = E \{ f^k(t) \} = t^{\beta(k)} b(k)
\]  

(12)

in which \( b(k) = \{ f^k(1) \} \) and \( \beta(k) = k\beta \) under simple scaling condition. Hence, the scaling behaviour of extreme rainfall can be examined by the power-form relationship between the \( k \)-order NCMs and the \( t \) durations. If extreme rainfall data exhibit the scaling properties, the log-linearity will be shown (Nguyen, 2004). From the log-linearity, the scaling exponent (\( \beta \)) can be determined as the slope on log scale.

In addition to the scaling properties in NCMs, for a simple scaling process, it can be shown that the statistical properties of the GEV distribution for two different time scales \( t \) and \( \lambda t \) are related as follows:

\[
k(\lambda t) = k(t)
\]  

(13)

\[
\alpha(\lambda t) = \lambda^\beta \alpha(t)
\]  

(14)

\[
\xi(\lambda t) = \lambda^\beta \xi(t)
\]  

(15)

\[
X(\lambda t) = \lambda^\beta X(t).
\]  

(16)

Based on these relationships it is possible to derive the statistical properties of short-duration (e.g., \( \lambda t = <1 \) day) extreme rainfalls using the properties of daily (\( t = 1 \) day) extreme rainfalls. The exponent is computed based on the scaling properties of the NCMs of extreme rainfalls for various durations. Consequently, the proposed scaling GEV can be used to derive the IDF relationships for AMPs for different durations.

Hence, given simple scaling properties in extreme rainfalls we are able to infer three parameter estimation methods from Equation 12 through 16 such as;

a) Scaling L-moment method (hereafter called SLmom).

After estimating extreme rainfall intensities for short-
duration ($\lambda t$) using Equation 16, apply L-moment method to the estimated AMPs for obtaining the GEV distribution parameters (see Hosking [1990]).

\[
\beta = \frac{\mu_1(\lambda t)}{\mu_1(t)}
\]  

(17)

b) Scaling One-NCM method (S1NCM). With simple scaling properties, the scaling factor $\lambda^\beta$ can be easily estimated by:

where $\mu_1(t)$ and $\mu_1(\lambda t)$ are the first order NCMs of the extreme rainfall for the durations $t$ and $\lambda t$, respectively. Once obtaining the scaling exponent with Equation 17,
the GEV parameters for the short-duration ($\lambda t$) can be estimated using Equations 13–15. With this method, the shape parameter for duration of $\lambda t$ is same as for duration of $t$.

c) Scaling Three-NCM method (S3NCM). If the data exhibits simple scaling, the first three NCs for the short-duration ($\lambda t$) can be estimated using Equation 12 and the scaling exponents. The estimated NCs will be used for estimating the parameters of GEV distribution for the specific durations ($\lambda t$) using Equations 5–7. Although there are three unknown variables and three equations, it is impossible to get directly analytical solutions due to an inherent gamma function. Numerical analysis, in this study, is conducted for calculating the shape parameter.

In this study, the applications of the above methods are presented for downscaling temporally extreme rainfall data from a long-time scale to short-time scales. Figure 1 shows the scheme of the presented downscaling methods. To evaluate the performances of them, the comparison study is carried out using historical data sets from two distinct at Dorval in Quebec (Canada) and Seoul (South Korea).

### 2.2 Modified bootstrapping for CIs for downscaled IDF curves

This section describes a modified bootstrap technique to determine CIs of the downscaled sub-daily extreme rainfall series. On the basis of a simple scaling behaviour of extreme rainfalls, the proposed method is carried out by the combination of bootstrapping technique and the S3NCM estimation method introduced in the previous section.

By bootstrapping technique, $n$ (e.g., here 1,000) sets of AMPs for daily duration $T$ are generated from observed/projected AMPs for the duration $T$. The next step is to estimate the $n$ sets of first three NCs from the $n$ sets of the resampled AMPs for the duration $T$. For constructing CIs, first the upper and lower limits of the NCs for duration $T$ at a given significant level (e.g., 95%) are estimated from the $n$ sets of the first three NCs. Based on simple scaling-invariance properties, the first NCs for shorter duration $t$ are derived from the determined upper and lower limits of the NCs for duration $T$. And then, parameters ($\xi$, $\alpha$, and $\kappa$) of the downscaled GEV distributions for each duration $t$ are calculated by Equations 5–7 and the upper and lower limits of the NCs for duration $t$. CIs at a given confidence level are constructed by computing rainfall quantiles from the estimated parameters and Equation 8. The computational procedures are summarized as the following (also shown in Figure 1);

a) Generate $n$ sets of AMPs for duration $T$ by bootstrapping technique with observed/projected daily AMPs, and then calculate the first three NCs for duration $T$;
b) Choose the upper and lower limits of NCs for duration $T$ at a given significant level;
c) Estimate the upper and lower limits of NCs for duration $T$ using Equation 12 and the scaling exponents;
d) Estimate parameters of the downscaled-GEV distributions for the upper and lower limits using the estimated first three NCs;
e) Calculate rainfall intensities for the upper and lower limits for each duration.

A major advantage of this procedure is that it requires a small number of data. Once getting the scaling exponents based on simple scaling properties of AMPs, the confidence intervals of IDF curves can be constructed using only daily AMPs and the proposed method.

### 3 DATA AND METHODOLOGY

For the feasibility and accuracy of the proposed temporal downscaling approaches and the estimation method for CIs of IDF curves, case studies were carried out using at-site AMP data available at two rain gauge stations located in two completely different climatic regions: Dorval station in cold-climate southern Quebec region in Canada and Seoul station in subtropical-climate South Korea. The annual precipitation at Dorval is 3,047 mm including an average snowfall of 2,262 mm and rainfall of 785 mm. With low seasonal variability, the precipitation at this station is spread evenly. On the contrary, Seoul has a humid subtropical climate described as humid continental with large variation of the precipitation through years by the monsoons. About 48.6% (668 mm) of the annual precipitation (1,373 mm) falls for only July and August. Historical at-site AMP series at the two stations for the 1961–1990 period were obtained from Canadian Weather Energy and Engineering Datasets (CWEEDS) and Korea Meteorological Administration (KMA). They were made up of nine durations (5-min, 10-min, 15-min, 30-min, 1-h, 2-h, 6-h, 12-h, and 1-day) and 16 durations (10-min, 20-min, 30-min, 40-min, 50-min, 1-h, 1.5-h, 2-h, 3-h, 4-h, 6-h, 9-h, 12-h, 15-h, 18-h, and 1-day) for Dorval and
Seoul stations, respectively. As shown in Figure 2, the numerical applications could be classified into the three sub-sections: the verification of parameter estimation methods, temporal downscaling including simple scaling test, and CIs for the estimated IDF curves.

4 | RESULTS AND DISCUSSIONS

4.1 | Scaling-GEV models

4.1.1 | Verification of parameter estimation methods

To demonstrate the alternative parameter estimation method based on the first three NCMs, the comparison study is carried out using the observed and estimated. Both the conventional estimation method based on L-moments and the proposed method are applied to the observed AMPs for parameterizing distributions for each duration. Quantiles are estimated with the exceedance probability and the determined sets of parameters (e.g., location, scale, and shape) of distributions for given durations. The estimated quantities from two different sets of parameters estimated by the L-moment based and NCM based approaches are evaluated by the numerical/graphical methods.

The numerical analyses using the relative-root-mean-square-error (RRMSE) are carried out to assess the two estimation methods. Because the root-mean-square-error (RMSE) gives more weights to larger values than those to smaller values, the values of RMSE are apt to be sensitive to the presence of outliers. On the other hands, RRMSE figures out each error proportion to each observed value (Tao et al., 2002). The RRMSE is expressed by:

$$ \text{RRMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \frac{X_{i}^{\text{Obs}} - X_{i}^{\text{Est}}}{X_{i}^{\text{Obs}}} \right)^2} $$

where $X_{i}^{\text{Obs}}$ and $X_{i}^{\text{Est}}$ are observed and estimated quantities, respectively. Like other error criteria, the smaller value of RRMSE indicates the better the fit is. Table 1 shows the values of RRMSEs by the estimation method using NCMs are very close to those by the method based on L-moments for both two stations.

In addition to numerical assessment, Figure 3 shows the quantile plots for duration 5-min and 1-h for two stations (Dorval and Seoul), respectively. Black, red, and blue dots in the figure represent observed, estimated quantities by L-moments and the first three NCMs, respectively. It is found that the estimated quantiles by the proposed NCMs based method are very close to not only those by the L-moment method, but also observed values.

The comparison results have indicated that the first three NCMs could have enough information to estimate GEV parameters and thereby can be used for the purpose of downscaling extreme rainfalls for shorter duration from observed/forecasted daily AMPs.

4.1.2 | Scaling properties of AMPs

The scaling properties of the at-site AMP series for all durations were examined using the first three NCMs of AM precipitations. According to Equation 12, the relationship between scaling property and NCMs can be plotted within log–log regime, that is, the log-linearity in this regime indicates the power law dependency (i.e., scaling) of NCMs with durations. Figure 4 shows the scaling relationships with respect to all duration. Scaling properties are observed at two different regimes of durations. The breakpoints at the stations are differently located of 30-min (Dorval) and 90-min (Seoul), respectively. The presence of breakpoint may imply the transition of precipitation dynamics from convective storm system to synoptic storm system (Vu et al., 2018).

Further, the linearity of the scaling exponent with the moment order as shown in Figure 5 supports the assumption that the extreme rainfall series considered can be described by a simple scaling model. Hence, for given locations, it is possible to determine the NCMs and the distribution of rainfall extremes for short durations (e.g., 30 or 90 min) using available rainfall data for longer time scales (e.g., 1 day) within the same scaling regime ($\beta$ is known).

**FIGURE 2** Scheme of the computational procedure
4.1.3 Compare the temporal downscaling methods

On the basis of the simple scaling properties, the proposed temporal downscaling models (i.e., SLmom, S1NCM, and S3NCM) are employed to estimate AM rainfalls corresponding to each shorter duration from observed daily AMPs (i.e., 24 h duration). The estimated quantiles by the three downscaling models are evaluated with the observed values by both numerical/graphical approaches. Tables 2 and 3 show the values of RRMSE for quantiles estimated by the three scaling models for Dorval and Seoul stations, respectively. Bold letter indicates the lowest value for each duration. That is, the

| Dorval (Canada) | Seoul (South Korea) |
|----------------|--------------------|
| **Duration**   | **L-mom** | **NCMs** | **Duration** | **L-mom** | **NCMs** |
| 5_min          | 0.095     | 0.093     | 10_min       | 0.098     | 0.096     |
| 10_min         | 0.095     | 0.094     | 20_min       | 0.104     | 0.102     |
| 15_min         | 0.103     | 0.102     | 30_min       | 0.103     | 0.103     |
| 30_min         | 0.112     | 0.110     | 40_min       | 0.112     | 0.111     |
| 60_min         | 0.114     | 0.111     | 50_min       | 0.116     | 0.116     |
| 2_h            | 0.095     | 0.093     | 1_h          | 0.120     | 0.119     |
| 6_h            | 0.074     | 0.071     | 1.5_h        | 0.115     | 0.113     |
| 12_h           | 0.084     | 0.082     | 2_h          | 0.114     | 0.112     |
| 24_h           | 0.071     | 0.069     | 3_h          | 0.110     | 0.108     |
|                |           |           | 4_h          | 0.109     | 0.108     |
|                |           |           | 6_h          | 0.106     | 0.104     |
|                |           |           | 9_h          | 0.102     | 0.099     |
|                |           |           | 12_h         | 0.107     | 0.104     |
|                |           |           | 15_h         | 0.116     | 0.113     |
|                |           |           | 18_h         | 0.120     | 0.117     |
|                |           |           | 24_h         | 0.133     | 0.132     |

*Note:* L-mom denotes the estimation method based on L-moments, and NCMs does the method using NCMs.

**Figure 3** Quantile plot comparing the observed 5-min and 1-h AMPs to the estimated values by L-moments method and NCM method for Dorval and Seoul, respectively. (a) 5-min quantile plot for Dorval, (b) 1-h quantile plot for Dorval, (c) 5-min quantile plot for Seoul, and (d) 1-h quantile plot for Seoul. The black dots denote the observed values, the orange line does the quantiles estimated by L-moment method, and the blue line represents those by the proposed NCM method.
model having a bold letter estimates quantiles that are the closest to observed quantiles. Numerical assessment results show that the Scaling-GEV distribution method using the first three NCMs give the best fit for on average sub-daily durations in both two stations. For the longer durations (e.g., over 12-h), the L-moment fitting method with simple scaling GEV distribution provides more accurate estimates. Regarding the shorter durations (e.g., less 12-h), the Scaling-GEV model using only the first one NCMs is the worst for Dorval station, while the SLmom model gives the worst estimates for Seoul station. Unlike Dorval station case, the S1NCM method provides a result that shows a better accuracy of about 5% or less than the S3NCM method for durations of 40-min, 50-min, and 1-h. Furthermore, Equation 13 indicates the S1NCM model has constant shape parameter over all durations. The GEV model with a fixed value of shape parameter may result in similar effect of Gumbel distribution (EV1), which has zero value of it, because the shape parameter governs the tail behaviour of the distribution. Hence, the results could imply that Gumbel distribution could be another candidate distribution for fitting extreme rainfalls for Seoul station. Nevertheless, Scaling-GEV with three NCMs gives best results over both stations.

For comparison purposes, the quantile plots and IDF curves for only two duration (i.e., 10-min and 1-h) are constructed using the proposed scaling GEV distribution

| Duration | Parameter estimation method |
|----------|-----------------------------|
| 5-min    | SLmom | S1NCM | S3NCM |
| 10-min   | 0.0119 | 0.0130 | 0.0075 |
| 15-min   | 0.0196 | 0.0249 | 0.0132 |
| 30-min   | 0.0224 | 0.0244 | 0.0139 |
| 1-h      | 0.0274 | 0.0290 | 0.0181 |
| 2-h      | 0.0219 | 0.0230 | 0.0201 |
| 6-h      | 0.0142 | 0.0040 | 0.0148 |
| 12-h     | **0.0122** | 0.0145 | 0.0134 |

Note: Bold letter indicates the lowest value amongst three models for each duration.
models. In Figures 6 and 7, black dots denote observed rainfall intensities, orange diamonds do the estimated AMPs by SLmom model, green triangles those by S1NCM model, and blue circles those by S3NCM model. The S3NCM model estimates very accurately extreme values having high value of the probability of exceedance, especially for Dorval station. For Seoul station, the S1NCM method appears to be sensitive to extreme quantities, so it estimates the closest values having the highest probability of exceedance to the observed. However, the S3NCM estimates the closest values to the observed over all durations. Hence, graphical analyses using quantile plot and IDF curves have indicated that the S3NCM could provide the most accurate estimates amongst the suggested parameter estimation models using the scale invariance properties for both stations, like numerical assessment results.

4.2 | Confidence intervals (CIs) for scaled AMPs

The proposed estimation method is used to construct CIs of AMPs for Dorval station (Quebec, Canada). In this numerical application, bootstrap technique is used to generate 1,000 sets of AMPs for duration 1-day using observed daily AMPs, and then the NCMs corresponding to the 95% significant level are estimated for constructing CIs of the first three NCMs (Figure 8a). Grey shade denotes the range of CIs that represent quantiles at 95% significant level. The selected NCMs are used to fit Scaling-GEV distributions for constructing the upper and lower boundaries. Through Figure 7b,c, it is found that the combination of scaling invariance properties of annual extreme rainfalls and the proposed CI method could provide robust CIs of rainfall estimates without the observed sub-daily AMPs. Based on the results, IDF curves corresponding to the 2-years and 100-years return periods are estimated (Figure 8d).

| Duration | Parameter estimation method |
|----------|-----------------------------|
|          | SLmom | S1NCM | S3NCM |
| 10-min   | 0.0307 | 0.0251 | 0.0152 |
| 20-min   | 0.0239 | 0.0181 | 0.0095 |
| 30-min   | 0.0262 | 0.0228 | 0.0171 |
| 40-min   | 0.0246 | 0.0217 | 0.0223 |
| 50-min   | 0.0256 | 0.0209 | 0.0220 |
| 1-h      | 0.0234 | 0.0174 | 0.0184 |
| 1.5-h    | 0.0337 | 0.0224 | 0.0133 |
| 2-h      | 0.0346 | 0.0215 | 0.0139 |
| 3-h      | 0.0317 | 0.0197 | 0.0162 |
| 4-h      | 0.0291 | 0.0197 | 0.0181 |
| 6-h      | 0.0344 | 0.0252 | 0.0239 |
| 9-h      | 0.0298 | 0.0219 | 0.0207 |
| 12-h     | 0.0254 | 0.0202 | 0.0194 |
| 15-h     | 0.0161 | 0.0169 | 0.0167 |
| 18-h     | 0.0115 | 0.0157 | 0.0155 |

Note: Bold letter indicates the lowest value amongst three models for each duration.
FIGURE 7  (a,c) Quantile plots and (b,d) IDF curves of AMPs for 10-min and 1-h durations estimated by the three simple scaling models for Seoul station. Black dots denote observed rainfall intensities, orange diamonds do the estimated AMPs by scaling L-mom model (SLmom), green triangles those by scaling one-NCM model (S1NCM), and blue circles those by scaling three-NCM model (S3NCM).

FIGURE 8  The 95% confidence intervals (CIs) of the first three (a) NCMs, (b) IDF curves for the 5-min, and (c) 1-h durations, and IDF curves and (d) CIs corresponding to 2 and 100 years return periods, respectively. Grey shade regions are the 95% CIs constructed by the proposed method, black lines are the estimated IDF curves by the Scaling-GEV method and the S3NCM estimation method, and blue dots represent observed quantiles.
5 | CONCLUSIONS

In this study, we evaluated the performances of three fitting methods for scaling GEV distributions in terms of the ability to downscale daily AMPs into those for the duration of sub-daily (e.g., 5-min, 10-min, and 1-h). Given observed daily AMPs and simple scaling properties of annual extreme values, we can construct daily and sub-daily IDF curves using Scaling L-moment method, Scaling One-NCM method, and Scaling Three-NCM method. These downscaling approaches are useful for assessing climate change impacts on extreme rainfalls and updating IDF curves for future condition in collaboration with outputs from Global Climate Models (GCMs) and/or Regional Climate Models (RCMs).

The fitting methods are used with annual extreme precipitation data from two distinct climate regions, Dorval, Quebec (Canada) and Seoul (South Korea), for the 1961–1990 period. It is found that the AMP series at Dorval and Seoul displayed a simple scaling behaviour within two different time intervals. Based on this scaling property, the fitting methods estimate the parameters of GEV distributions for shorter durations. Overall, the scaling GEV distribution model with the three-NCM fitting method has been shown to be able to provide the most accurate estimates of sub-daily AMPs from daily AMPs. Scaling One-NCM method assumes a constant shape parameter regardless durations, while Scaling Three-NCM method calculates each shape parameter for each GEV distribution. Because a shape parameter is determined by skewness of distribution, the numerical comparison results could indicate that a shape parameter of GEV distribution plays an important role in capturing extreme values corresponding to probability of tail.

Two parameter estimation methods, which are based on L-moments and the first three NCMs, were applied to estimate sets of distribution parameters of the observed data sets from Dorval and Seoul stations. From the comparison results, it could be concluded that the alternative estimation method can be used for linking daily AMPs to sub-daily AMPs because the first three NCMs have enough information to figure out GEV distributions.

Furthermore, this study suggests a new estimation method for confidence intervals of extreme rainfall series. Although the CIs were constructed by only daily AMPs and the simple scaling properties, the observed sub-daily AMPs are generally within the 95% CI estimated. This CI estimation method may be useful for the uncertainty study for climate change impact studies and with sparse data available regions.

REFERENCES

Aronica, G. and Freni, G. (2005) Estimation of sub-hourly DDF curves using scaling properties of hourly and sub-hourly data at partially gauged site. Atmospheric Research, 77, 114–123.

Bara, M., Kohnová, S., Gaál, L., Szológy, J. and Hlavcová, K. (2009) Estimation of IDF curves of extreme rainfall by simple scaling in Slovakia. Contributions to Geophysics and Geodesy, 39, 187–206.

Blanchet, J., Ceresetti, D., Molinié, G. and Creutin, J.-D. (2016) A regional GEV scale-invariant framework for intensity–duration–frequency analysis. Journal of Hydrology, 540, 82–95.

Bougadis, J. and Adamowski, K. (2006) Scaling model of a rainfall intensity-duration-frequency relationship. Hydrological Processes, 20, 3747–3757.

Buishand, T. (1989) Statistics of extremes in climatology. Statistica Neerlandica, 43, 1–30.

Burlando, P. and Rosso, R. (1996) Scaling and multiscale models of depth-duration-frequency curves for storm precipitation. Journal of Hydrology, 187, 45–64.

Burn, D.H. (2003) The use of resampling for estimating confidence intervals for single site and pooled frequency analysis/Utilisation d’un rééchantillonnage pour l’estimation des intervalles de confiance lors d’analyses fréquentielles monomulti-site. Hydrological Sciences Journal, 48, 25–38.

Chernick, M.R. (2011) Bootstrap Methods: A Guide for Practitioners and Researchers. Hoboken, NJ: John Wiley & Sons.

Choi, J., Lee, O., Jang, J., Jang, S. and Kim, S. (2019) Future intensity–depth–frequency curves estimation in Korea under representative concentration pathway scenarios of fifth assessment report using scale-invariance method. International Journal of Climatology, 39, 887–900.

Fadhel, S., Rico-Ramirez, M.A. and Han, D. (2017) Uncertainty of intensity–duration–frequency(IDF) curves due to varied climate baseline periods. Journal of Hydrology, 547, 600–612.

Feder, J. (2013) Fractals. New York: Springer Science & Business Media.

Hershfield, D. M. 1961. Rainfall frequency atlas of the United States. Technical Paper: 40.

Hogg, W.D. and Carr, D. (1985) Rainfall Frequency Atlas for Canada. Ottawa: Environment Canada, Atmospheric Environment Service.

Hosking, J.R. (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society: Series B (Methodological), 52, 105–124.

Kite, G. (1974) Assessment of the Assumption of Normality in Computation of Confidence Limits. Ottawa, Canada: Department of Environment Canada, Inland Waters Directorate, Water Research Bureau.

Kotegoda, T. and Rosso, R. (1997) Statistics, Probability and Reliability for Civil and Environmental Engineers. New York: McGraw-Hill Publishing Company.
Koutsoyiannis, D. (2004) Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique. *Hydrological Sciences Journal*, 49(4), 575–590.

Lu, Y., Qin, X. and Mandapaka, P. (2015) A combined weather generator and K-nearest-neighbour approach for assessing climate change impact on regional rainfall extremes. *International Journal of Climatology*, 35, 4493–4508.

Mailhot, A., Duchesne, S., Caya, D. and Talbot, G. (2007) Assessment of future change in intensity–duration–frequency (IDF) curves for southern Quebec using the Canadian regional climate model (CRCM). *Journal of Hydrology*, 347, 197–210.

Menabde, M., Seed, A. and Pegram, G. (1999) A simple scaling model for extreme rainfall. *Water Resources Research*, 35, 335–339.

NERC. (1975) Flood studies report. *Flood Studies Report*, Wallingford, UK: Institute of Hydrology.

Nguyen, T.-D. (2004) Regional estimation of extreme rainfall events. Doctoral Thesis. Department of Civil Engineering and Applied Mechanics, McGill University, Montreal (Quebec), Canada.

Nguyen, V., Desramaut, N. & Nguyen, T. 2008 Estimation of urban design storms in consideration of GCM-based climate change scenarios. *Proceedings of International Conference on Water & Urban Development Paradigms: Towards an integration of engineering, design and management approaches*, Leuven, 2008. pp. 347–356.

Nguyen, V., Nguyen, T.-D. and Ashkar, F. (2002) Regional frequency analysis of extreme rainfall. *Water Science and Technology*, 45, 75–81.

Nhat, L. M., Tachikawa, Y., Sayama, T. & Takara, K. 2008. Estimation of Sub-Hourly and Hourly IDF Curves Using Scaling Properties of Rainfall at Gauged Site in Asian Pacific Region.

Pandey, G.R. (1997) Regional Estimation of Floods and Rainfalls for Ungauged Sites. *Journal of Hydrologic Engineering*, 21(12), 04016044.

Reed, D.W., Faulkner, D., Robson, A., Houghton-Carr, H. and Bayliss, A. (1999) *Flood Estimation Handbook (Procedures for Flood Frequency Estimation)*. Wallingford. Oxfordshire. UK: Institute of Hydrology.

Schaefer, M. (1990) Regional analyses of precipitation annual maxima in Washington state. *Water Resources Research*, 26, 119–131.

Tao, D., Nguyen, V. & Bourque, A. 2002 On selection of probability distributions for representing extreme precipitations in Southern Quebec. *Annual Conference of the Canadian Society for Civil Engineering*, 2002. pp. 1–8.

Vu, M., Raghavan, S., Liu, J. and Liong, S.Y. (2018) Constructing short-duration IDF curves using coupled dynamical–statistical approach to assess climate change impacts. *International Journal of Climatology*, 38, 2662–2671.

Wilks, D.S. (1993) Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. *Water Resources Research*, 29, 3543–3549.

Yeo, M.-H. 2013. *Statistical modeling of precipitation processes for gaged and ungaged sites in the context of climate change*. Thesis, McGill University Libraries.

Yu, P.-S., Yang, T.-C. and Lin, C.-S. (2004) Regional rainfall intensity formulas based on scaling property of rainfall. *Journal of Hydrology*, 295, 108–123.

Zalina, M.D., Desa, M.N.M., Nguyen, V. and Kassim, A.H.M. (2002) Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science And Technology*, 45, 63–68.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Yeo M-H, Nguyen V-T, Kpodonu TA. Characterizing extreme rainfalls and constructing confidence intervals for IDF curves using Scaling-GEV distribution model. *Int J Climatol*. 2021;41:456–468. [https://doi.org/10.1002/joc.6631](https://doi.org/10.1002/joc.6631)