New Method of Deriving IDF Curve & Analyze Rainfall Intensity

Kanghee Lee

Abstract

This application derives the IDF (Intensity-Duration-Frequency) curve if the user put the historical rainfall data. Compare to the previous methods, it applies the locally weighted linear regression, so it comes to be simpler than the previous methods.

Introduction

Because of the climate change the importance of water management is regarded more important than before. For example, California is in trouble with drought and Texas got a huge damage from the flood. Therefore, to handle this problem, Hydrologist predict the abnormal flood or drought from the past rainfall data and design the hydraulic sturctures to prevent the natural disaster occur by the rainfall.

By analyzing historic rainfall data, rainfall frequency can be derived and get the data such as 100 year frequency rainfall or drought. Hydrologists derive the IDF(Intensity, Duration, Frequency) curve each time period analyzed rainfall data, these data decrease by the log scale. Nowadays, they use several methods to derive the IDF curve, and these methods needs to derive parameters for each region.

This application analyzes and derives the rainfall frequency if the user puts the historic rainfall data. Moreover, this application derives the IDF curve by applying the linear regression algorithm, so it can be applied regardless of the region, without finding parameters for each region.

Related Work

There are several empirical equation to derive rainfall intensity. Nowadays, to get more accurate rainfall intensity, following rainfall intensity formula is used.

$$I(T,t) = \frac{a + b + \ln T/t^n}{c + d \times \ln \sqrt{T}/t + \sqrt{T}}$$

I is the rainfall intensity, *T* is return period(year), *t* is duration(min).

a, b, c, d and n are the parameters that are determined by region. Other empirical equations also need the parameter for each region

Data Set

  The figure below is the plot of the historic rainfall data at Suwon and Icheon city (South Korea) weather station (1973~2014, 1hour~24hours). To apply the hydrologic year system, the 2015 rainfall data is not used. This rainfall frequency and the IDF curve will be derived by this application. Suwon city and Icheon city(South Korea) data will be analyzed too for checking that this application can be used comprehensively for other regions.

<Figure 1. Suwon & Icheon City rainfall dataset>
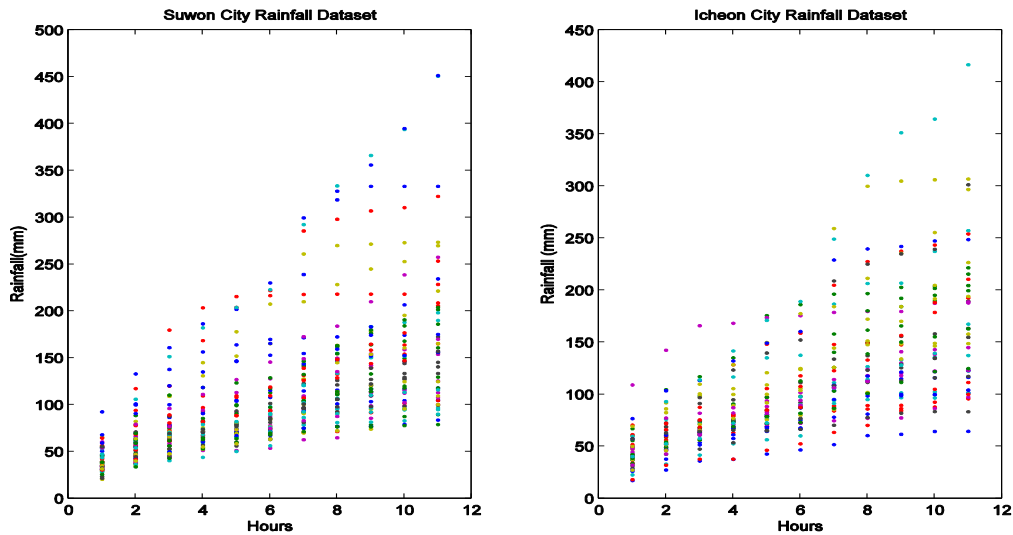
  The following figure shows the 1 hour duration rainfall at Suwon & Icheon station.

<Figure 2. Suwon & Icheon City 1hour duration rainfall>
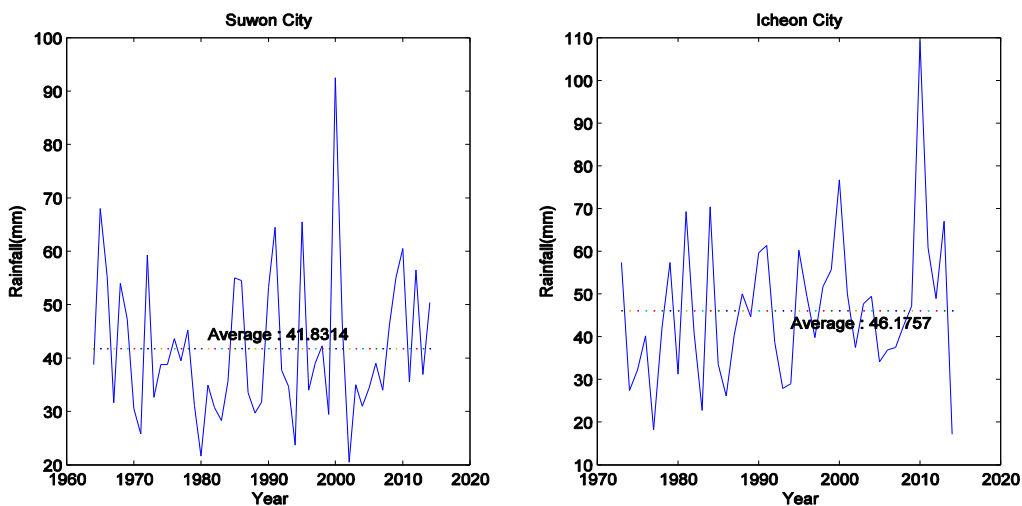
  According to the figure above, the tendency of the rainfall have the skewness, that the data is not normal distribution, there are not many data close to the average. Therefore, to get the rainfall frequency, several methods are used to calibrate the skewness. This application the most recommended method ; GUMBEL method are used. Both dataset is from K-water(Korea Water Resources Corporation).

Methods

Gumbel method requires specific moment values to calibrate skewness.
Probability weighted moment

$$M_{1,0,0} = E[X] = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$M_{1,1,0} = E[XF(x)] = \frac{1}{n}\sum_{i=1}^{n} \frac{(i-1)}{(n-1)} x_i$$

$$M_{1,2,0} = E[XF^2(x)] = \frac{1}{n}\sum_{i=1}^{n} \frac{(i-1)(i-2)}{(n-1)(n-2)} x_i$$

x_i is the (i)th value of the ascending rainfall data
L-moment

$$L_1 = E[X] = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$L_2 = \frac{1}{2} E[X_{(1:2)} - X_{(2:2)}]$$

$$L_3 = \frac{1}{3} E[X_{(1:3)} - 2X_{(2:3)} + X_{(3:3)}]$$

Gumbel distribution method. Cumulative Probability Density Function is

$$F(x) = \exp\left[-\exp\left(-\frac{x - x_0}{\alpha}\right)\right] = 1 - \frac{1}{T}$$

$$\alpha(scale\ variable) = L_2 / \ln 2$$

$$x_0(positional\ argument) = L_1 - 0.05772\alpha$$

Because of the skewness the positional argument and scale variable is applied. By the equation above, rainfall by time period derived as

$$x = x_0 - \alpha \ln\left[\ln\left(\frac{T}{T-1}\right)\right]$$

By the result of calibration and analysis of historical rainfall data, we can receive the rainfall intensity of each time period. By this result IDF (Intensity-Duration-Frequency) 'plot' can be drawn. The previous methods of deriving IDF curve estimate the curve by the rainfall intensity per frequency, duration and the parameters for each region. This application applies Locally weighted linear regression.

First of all to Maximized log likelihood

$$l(\theta) = \log L(\theta)$$

$$= \log \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right)$$

$$= \sum_{i=1}^{m} \log \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right)$$

$$= m \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\sigma^2} \cdot \frac{1}{2} \sum_{i=1}^{m} (y^{(i)} - \theta^T x^{(i)})^2$$

It is same as minimizing

$$l(\theta) = \frac{1}{2} \sum_{i=0}^{m} (y^{(i)} - \theta^T x^{(i)})^2$$

Standard choice for the weights is applied for this application.

$$w^{(i)} = exp\left(-\frac{(x^{(i)} - x)^2}{2\tau^2}\right)$$

For this application the τ value needs to be derived by the dataset of first region and apply this value at the other region to prove that it is applicable to other regions.
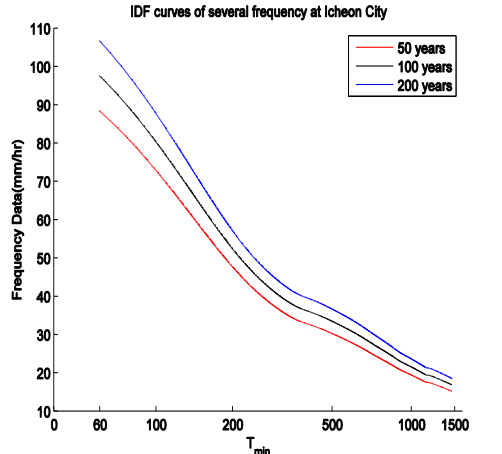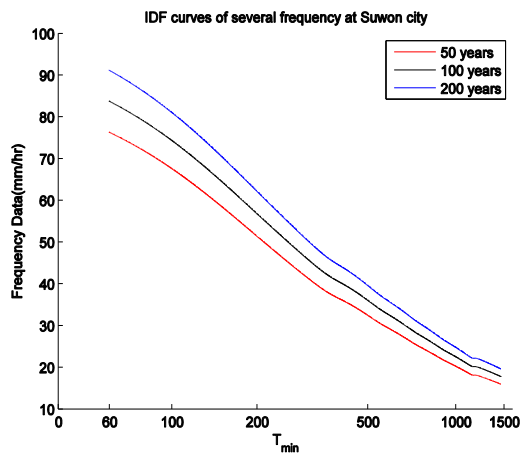
Result & Discussion

To get the weights,

Plot and the linear regression of data several tau values (Suwon City)
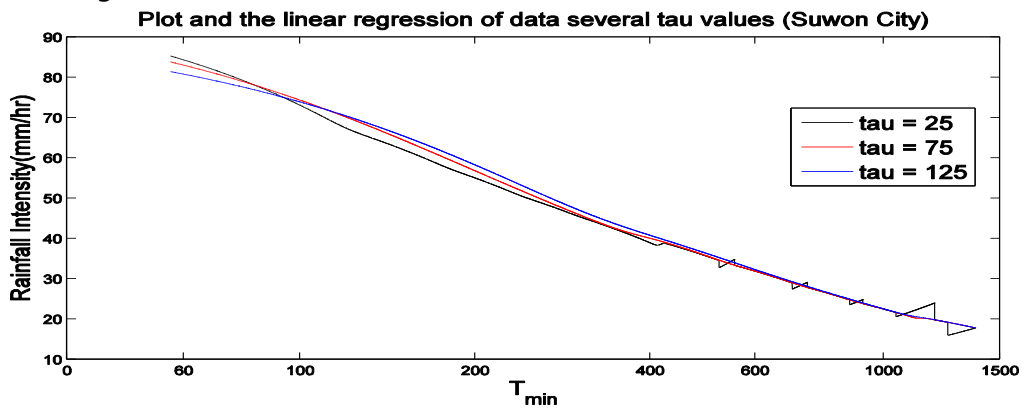
By the result of the above from Suwon city data the τ value is estimated as 75, and it is also applicable for Icheon City rainfall data as shown as below.

Plot and the linear regression of data several tau values(Icheon City)
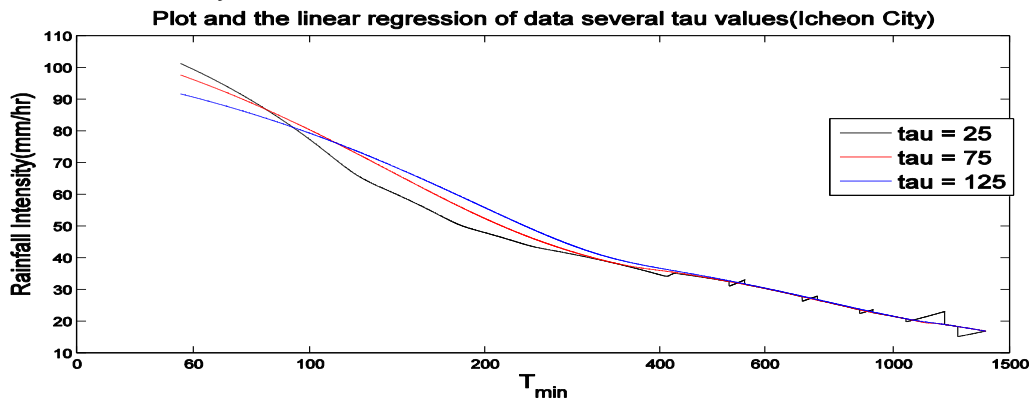
By applying the τ as derived, the above figures are the derived IDF curves of several frequencies, by that τ value.

Previous methods need several parameters to derive the IDF curve and rainfall intensity for each region. Moreover, more accurate method needs more complicate and more various parameters. However, this application is independent from these parameters so it is applicable at various regions. Therefore, to derive the IDF by this application, only historical rainfall data is needed.

Future Work

The bandwidth parameter needs to be more accurate by comparing the results of other various regions rainfall data to be more applicable for various regions. Moreover, it is extreme climate regions such as desert, polar, tropical regions rainfall data is needed to compare and check that is applicable for those climates.

References

Jea Su Lee, Hydrology, GUMI 2012

Ng, A. CS 229 lecture notes. Part I – Linear Regression