



Stochastic disaggregation of daily rainfall based on K-Nearest neighbor resampling method

Park, HeeSeong^a · Chung, GunHui^{b*}

^aHydro Science and Engineering Research Institute, Korea Institute of Civil Engineering and Building Technology, Goyang, Gyeonggi-do, 10223, Korea

^bDepartment of Civil Engineering, Hoseo University, Asan, Chungcheongnam-do, 31499, Korea

Paper number: 15-089

Received: 26 October 2015; Revised: 12 February 2016 / 17 February 2016; Accepted: 17 February 2016

Abstract

As the infrastructures and populations are condensed in the mega city, urban flood management becomes very important due to the severe loss of lives and properties. For the more accurate calculation of runoff from the urban catchment, hourly or even minute rainfall data have been utilized. However, the time steps of the measured or forecasted data under climate change scenarios are longer than hourly, which causes the difficulty on the application. In this study, daily rainfall data was disaggregated into hourly using the stochastic method. Based on the historical hourly precipitation data, Gram Schmidt orthonormalization process and K-Nearest Neighbor Resampling (KNNR) method were applied to disaggregate daily precipitation into hourly. This method was originally developed to disaggregate yearly runoff data into monthly. Precipitation data has smaller probability density than runoff data, therefore, rainfall patterns considering the previous and next days were proposed as 7 different types. Disaggregated rainfall was resampled from the only same rainfall patterns to improve applicability. The proposed method was applied rainfall data observed at Seoul weather station where has 52 years hourly rainfall data and the disaggregated hourly data were compared to the measured data. The proposed method might be applied to disaggregate the climate change scenarios.

Keywords: Gram-Schmidt orthonormalization process; K-Nearest Neighbor Resampling (KNNR) method; Rainfall pattern; Urban runoff; Stochastic disaggregation of rainfall data

K번째 최근접 표본 재추출 방법에 의한 일 강우량의 추계학적 분해에 대한 연구

박희성^a · 정건희^{b*}

^a한국건설기술연구원 수자원하천연구소, ^b호서대학교 건축토목환경공학부 토목공학전공

요 지

산업의 발전에 따라 기반시설 및 인구 등이 대도시에서 밀집되어, 도시홍수방어는 인명피해 뿐만 아니라 재산피해 저감 차원에서도 매우 중요한 문제가 되었다. 요즘은 이러한 도시유역의 유출해석을 보다 정확하게 하기 위해 시강우나 분단위의 강우자료를 활용하고 있다. 하지만 기후변화 시나리오와 같은 미래 강우시나리오는 현재 일단위 수준으로 제공되므로 미래 강우에 대한 확률빈도 해석에 제한이 있다. 이에 본 연구에서는 추계학적 기법을 이용해 일강우 자료를 시강우 자료로 분해하고자 하였다. 일자료를 시자료로 분해하기 위해 과거 시강우 자료를 기반으로 Gram Schmidt 변환과 K개의 최근접 표본 중 하나를 재추출하는 비모수적인 기법(KNNR)을 적용하였다. 이 방법은 연유출량을 월유출량으로 분해하기 위해 개발된 것이다. 하지만 강우자료는 유출량 자료와 달리 확률밀도가 작아 일강우를 시강우로 분해하는 데 직접 적용하는 경우 결과가 실제와 유사한 통계 패턴을 갖는다고 보기 어려웠다. 이를 보완하기 위해 본 연구에서는 분해하고자 하는 일자의 전일과 후일을 포함한 3일 강우패턴을 7개로 구분하고 동일 패턴을 가진 자료들만 분해에 이용하도록 하여 강우자료에 대한 적용성을 높였다. 과거 52년간의 서울기상관측소 시강우 자료를 이용하여 강우자료의 분해에 대한 결과를 분석한 결과, 분해된 시강우 자료가 관측된 시강우자료와 통계적으로 매우 유사한 것을 확인하였다. 향후 기후변화자료의 시강우 분해 등에 활용하여 보다 정확한 도시유출에 대한 빈도해석 등에 적용할 수 있을 것으로 판단된다.

핵심용어: 강우자료 추계학적 분해, 도시유출, Gram Schmidt 변환, 강우 패턴, K-NN 표본 재추출 방법(KNNR)

*Corresponding Author. Tel: +82-41-540-5794
E-mail: gunhui@hoseo.edu (G.H. Chung)

1. 서론

인구밀도가 높은 도시지역의 강우를 배제하기 위해 우수관을 이용한 유출 시스템이 주로 이용된다. 우수관을 따라 하천 혹은 우수지 등으로 유출이 발생하는 유역의 구분을 배수분구라고 하며, 각 배수분구의 도달시간은 대부분 1~2 시간 이내이다. 이런 경우 도시홍수유출을 해석하기 위해서는 1시간 이하의 짧은 시간단위의 자료를 사용하여야 한다. 게다가 이러한 유출에 대해 통계적으로 유효한 빈도분석을 위해서는 30년 이상 오랜 기간의 관측자료가 필요하다. 특히 기후변화 자료에 대해 빈도분석을 수행하기 위해서는 시 단위 이하로 자료 분해가 필요하다. 이러한 경우, 유량자료에 대해서는 과거의 일자료를 분해하여 장기간의 시자료를 확보하는 방법이 오래 전부터 연구되고 있으나(Matalas, 1967; Mandelbrot and Wallis, 1968; 1969a; 1969b; 1969c; Rodriguez-Itube et al., 1972; Valencia and Schaake, 1973) 강우자료에 대해서는 그 적용성이 확인되지 않은 경우가 많다.

유량자료의 시간적 상세화 연구들을 살펴보면, Valencia and Schaake (1973)는 하천유량자료를 이용하여 연자료를 월자료로 분해하는 방법을 제안한 바 있으며, 해당 연구에서는 기존에 Matalas (1967), Mandelbrot and Wallis (1968, 1969a, 1969b, 1969c), Rodriguez-Itube et al. (1972) 등에서 제안한 부분잡음근사법(Fractional noise approximation)을 이용하여 추계학적으로 접근하였다. Valencia and Schaake (1973)는 계절적 변동성을 가지는 월자료의 합이 연자료와 같아지도록 fractional noise approximation 방법을 적용하여 연자료를 월자료로 분해하는 방법을 제안하였다. 그러나 이러한 방법들은 차원이 커질수록 적용이 매우 복잡해져 비교적 적용이 쉬운 비모수적 방법이 다양한 연구자들에 의해 시도되었다.

대표적인 비모수적인 방법으로 kernel density estimation (Tarboton et al., 1998) 방법과 Bootstrap 기법을 응용한 K번째 최근법 표본 재추출 방법(K-Nearest Neighbor Resampling Method; KNNR Method) (Kumar et al., 2000, Prairie et al., 2007)이 다수 적용되었다. 이들 중 kernel density estimator 방법은 고차원 문제에서는 매우 효율이 낮아지는 단점이 있다 (Sharma and O'Neil, 2002).

KNNR 방법은 모수에 대한 분포형 가정이 필요하지 않은 대표적인 비모수적 방법의 하나로, 대체로 독립동일분포인 자료들을 복원추출하는 방법을 이용한다. 그러나 Lall and Sharma (1996)는 상관성이 있는 비독립적인 자료들에 대해서도 적용이 가능한 KNNR 방법을 제안하고 이를 이용해 하

천유량 시계열 자료를 발생시킨 바 있다.

Kumar et al. (2000)은 월유량자료를 일유량자료로 분해하기 위해 KNNR 방법을 최적화 방법과 결합하여 적용하였다. 그러나 최적화 방법의 계산시간이 매우 길어 효율적이지 않았다. 또한 Prairie et al. (2007)도 연유량자료를 월유량자료로 분해하기 위해 KNNR 방법을 적용하였으나, 기존 자료의 월자료들 사이의 크기 차이에 따라 분해하는 대상 연유량자료를 나누므로, 분해하기 위한 연유량자료가 추출된 기존 연자료보다 작고, 월자료들 사이의 차이가 클 경우 간혹 분해된 결과가 0보다 작은 값이 나올 수 있다. Prairie et al. (2007)은 0보다 작은 값의 출현확률이 0.4% 이하이므로 크게 문제되지 않는다고 지적한 바 있다. 그 후, Lee et al. (2010)은 비모수적 세분화 모델인 KNNR 방법을 유전자 알고리즘과 결합하여 Colorado강의 연 하천유량을 월유량으로 분해하는데 적용한 바 있다.

강우자료의 경우 월자료의 일자료 분해에 대한 연구가 주를 이루고 있으며, 국내의 경우 기상청에서 이미 일자료로 분해된 기후변화 시나리오를 제공하고 있어 이에 대한 공간적인 상세화가 주로 연구되고 있는 상황이다.

국내에서 수행된 강우자료의 시간적인 분해 연구를 살펴보면, Kim et al. (2008)이 구형펄스모형과 분해기법을 적용하여 시강우자료로 변환한 뒤 I-D-F 곡선을 작성하여, 현재와 미래의 확률강우량의 변화를 평가하였다. 하지만 시강우자료에 대한 적정성 검토를 수행하지 않아 자료에 내재된 불확실성의 정도를 판단하기 힘든 한계점이 있다. 또한 Kyoung et al. (2008)은 강우자료에 대해 카오스적 특성을 보이는지 살펴보고 이를 이용해 일 강우자료에 대해 시간적 분해를 실시한 바 있다. 또한 Kim et al. (2014)는 비정상 은닉 마코프 모형을 개발하고 GCM 모의 결과를 통계학적으로 상세화하여 계절강수량을 일강수량으로 상세화하였다.

Lee and Jeong (2014)은 일강우자료를 시강우자료로 상세화하기 위해 KNNR 방법과 유전자 알고리즘을 결합한 방법을 적용하며, 시간단위의 통계학적 특성과 하루 동안의 시간별 통계적 패턴(낮의 주기)을 재현해 내는 비모수적 상세화 기법 모델을 개발하였다. 그 후, Lee et al. (2014)는 Lee and Jeong (2014)에서 제안한 방법을 부산과 경남지방의 기후변화 대표농도경로(RCP) 시나리오에 적용하였다.

이러한 기법들 중에서 KNNR 방법은 비모수적인 bootstrapping 기법에 바탕을 두고 있어 비교적 단순한 알고리즘으로 자료를 재현할 수 있다는 특징이 있어 다양한 분야에 활용되고 있다. 하지만 계절적인 경향성이 뚜렷한 월 유량의 추계학적 생성이나 연 유량의 월 유량 분해 등에 의해 생성되는 월

유량 자료는 상당히 원본과 근사한 통계적 경향성을 갖는 것으로 보이지만 일 강우자료를 특별한 경향성을 보이지 않는 시 강우자료로 분해하는 경우에는 분해된 자료에서 원본 자료의 통계적 경향성이 유지되는지 확인할 필요가 있다. 이에 본 연구에서는 Prairie et al. (2007)이 연 유량 자료를 월 유량 자료로 분해하는데 사용한 KNNR 방법을 일 강우자료를 시 강우자료로 분해하는데 적용할 수 있도록 보완하는 것을 목적으로 하였다. 이를 위하여 Prairie et al. (2007)의 방법에 따라 분해된 시 강우자료와 원본 시 강우 자료의 통계적 특성들이 유지되는지 평가하였으며, 경향성의 부재에서 발생하는 통계적 왜곡을 보완하기 위해 분해 대상일의 전일과 후일에 대한 3일 간의 강우패턴을 고려해 시 강우의 경향성이 고려되게 함으로써 기존의 방법을 개선하고 개선된 방법에 의해 분해된 시 강우자료와 원본 시 강우자료의 통계적 특성을 비교하여 개선효과를 검증하였다.

2. 이론적 배경 K-NN 표본 재추출 방법을 이용한 시계열 자료의 분해 방법

먼저 시간 분해에 대한 기본 개념은 오래전 Valencia and Schaake (1973)의 제안을 토대로 하고 있다. 분해된 강우시계열 값들이 총 d 개라고 했을 때, 분해된 자료들 (X_1, X_2, \dots, X_d) 의 합이 원 자료(Z)와 같아야 한다.

$$X_1 + X_2 + \dots + X_d = Z \tag{1}$$

여기서, d =분해된 시강우 벡터의 개수(차원), Z =일강우량이다.

분해된 시강우량의 합이 일강우량과 같아야 한다는 제약 조건을 이용하면 PDF $f(X|Z)$ 의 조건부 확률밀도함수를 Eq. (2)와 같이 만들 수 있다.

$$f(X|Z) = f(X, Z) / \int f(X, Z) dX \tag{2}$$

Eq. (2)의 분자(numerator)를 계산하기 위해서 $(d+1)$ 차원의 결합확률밀도함수 $f(X, Z)$ 가 필요한데 일강우량을 시강우자료로 분해하기 위해서는 25개의 확률밀도함수를 구하거나 커널함수 등을 이용해 확률함수를 조합할 필요가 있다. 이러한 방법은 매우 복잡하며, 실제 확률분포를 정확히 반영할 수 있는 확률함수를 추정한다는 것이 가능할 것인지에 대한 의문

이 여전히 남아 있다. 확률변수에 대한 모수에 대한 추정이 불가능하거나 곤란한 경우 모수에 대한 추정 자체를 생략하고 직접 확률변수를 모의 발생시키는 부트스트래핑과 같은 방법이 사용되기도 한다. KNNR 방법은 부트스트래핑과 같이 확률함수의 추정과 조합을 생략하고 직접 확률변수를 모의 발생시키는 방법의 하나이다.

KNNR 방법을 이용한 연유량자료의 월유량 분해에 대한 수행 절차는 Prairie et al. (2007)에 자세히 설명되어 있다. 본 연구에서는 일강우량자료를 시강우량으로 분해하기 위해 먼저 과거 n 개의 시강우자료 관측치의 집합을 다음과 같이 매트릭스 $X(d \times n)$ 형태로 나타내었다.

$$X = \begin{matrix} & \text{day1} & \text{day2} & \text{day3} & \dots & \text{dayn} \\ \begin{matrix} 1:00 \\ 2:00 \\ 3:00 \\ \dots \\ 24:00 \end{matrix} & \begin{matrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{matrix} & \begin{matrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{matrix} & \begin{matrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{matrix} & & \begin{matrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{matrix} \end{matrix} \tag{3}$$

여기서 d 는 분해하고자 하는 자료의 수, 즉 24이다. 또한, 과거 n 개의 시강우자료의 합인 일강우자료 벡터를 $Z(1 \times n)$ 라고 한다. Z 값을 분해하기 위해서는 Gram Schmidt 변환과 그 역변환을 이용한다. Gram Schmidt 변환은 X 의 한 열벡터를 Eq. (3)에 주어진 평면에 직각으로 회전 변환하는 과정으로 회전 매트릭스 R 을 구하는 방법은 Tarboton et al. (1998)에 자세히 설명되어 있다. X 의 열벡터들을 개별적으로 Eq. (4)에 의해 변환한다.

$$Y^i = R \cdot X^i \tag{4}$$

변환된 Y^i 벡터를 모으면 $(d \times n)$ 의 Y 매트릭스를 얻을 수 있다. 이때 회전 매트릭스 R 은 변환된 매트릭스 Y 의 마지막 행을 $Y_d = Z / \sqrt{d}$ 로 바꾼다. 따라서 적당한 Y 벡터를 선택해 Y_d 를 모의 발생시켜 기존 값을 대치하고 이를 역변환 함으로써 분해된 시강우를 얻을 수 있다. 이때 집합 Y 에서 사용하고자 하는 벡터를 선택하는 데는 Y_d 값을 변수로 하는 최근접 표본 재추출 방법을 이용한다. 즉 일강우자료를 z^* 라면, $z' = z^* / \sqrt{d}$ 를 계산하고 계산된 z' 와 변환된 매트릭스 Y 의 마지막 행을 Y_d 를 비교하여 k 개의 가장 가까운 값들에 대해 가까운 순서대로 가중치를 부여하여 확률적으로 그 중 하나를 선택한다.

이때 가중치는($W(k)$)는 다음 함수로 구한다.

$$W(k) = \frac{1}{k \sum_{i=1}^K \frac{1}{i}} \quad k = 1, 2, \dots, K \quad (5)$$

이 가중치함수는 가까운 이웃에게는 큰 가중치를 먼 이웃에게는 작은 가중치를 부여하는데, Lall and Sharma (1996)는 제안된 가중치가 균등분포 가중치에 비해 결과의 왜곡이 적고, 분산도 작아서 균등분포를 이용한 가중치에 비해 적절하다고 판단하였다. Eq. (5)에서 근접한 점들 중 몇 개(K)를 추출 대상으로 해야 하는지를 결정해야하는데 이는 휴리스틱하게 $K = \sqrt{N}$ 으로 결정해도 무방한 것으로 알려져 있다(Lall and Sharma, 1996). 이때 N은 샘플자료의 수이다.

선택된 K개의 이웃들 중에서 가중치를 고려하여 1개의 일강우 자료를 무작위 추출한다. 추출된 일강우 자료와 그에 해당하는 시강우 자료의 분포는 Y매트릭스 중 1개의 열벡터이며, 이를 $Y^*(d \times 1)$ 라고 하면, 변환 매트릭스 R을 이용하여 시강우 자료로 역변환을 하여 분해를 완성한다.

$$x^* = R^T Y^* \quad (6)$$

여기서 $x^*(d \times 1)$ 는 분해된 시강우자료 벡터이다.

3. 강우패턴을 고려한 분해 방법

기준에 제안된 시계열 자료의 분해는 Valencia and Schaake (1967)에서 Prairie et al. (2007)에 이르기까지 대부분 하천유량자료의 분해에 적용되었다. 반면 본 연구에서는 이를 강우자료에 적용하였는데 강우자료는 하천유량자료에 비해 확률함수의 밀도가 작기 때문에 이를 보완하기 위하여 3일 강우패턴을 이용하였다. 사용한 3일 강우패턴은 Ormsbee (1989)와 유사한데, Ormsbee (1989)는 시강우 자료를 20분 단위로 세분하기 위해 6개 강우패턴을 사용한 반면, 본 연구에서는 과거 일강우자료에 대해 선택된 일자의 전일과 후일 강우량을 비교하여 7개의 패턴으로 구분하고, 그 패턴에 해당하는 자료만을 대상으로 함으로써 확률함수의 밀도를 높일 수 있었다.

Ormsbee (1989)는 4개의 강우패턴을 사용한 Liou (1970)의 방법을 발전시켜 강우의 발생패턴을 6개로 나누어 제안하고, 패턴에 따라 시강우를 20분 단위로 분해하는 방법을 제안하였다. 이때 제안된 6개의 패턴은 Fig. 1(a)~(f)과 같은 형태를 갖고 있다. 패턴의 가운데 위치한 강우량 발생시간은 현재 시간 t라고 하면, t-1시간에서 t+1시간까지 강우가 감소하는 패턴(a, c, e)과 증가하는 패턴(b, d, f)으로 나뉜다.

본 연구에서는 Ormsbee (1989)가 제안한 6개의 패턴을 일강우자료에 적용하고 Fig. 1(g)의 패턴을 추가하여 총 7개의 패턴을 고려하였다. Type VII는 분해를 하려는 일자의 하루 전과 후의 강우가 같은 경우이다. 새롭게 제안된 Type VII 패턴을 포함하여 과거 52년간의 서울 관측소의 자료를 분리해 본 결과, Fig. 2에서 보이는 것과 같이 Type VII가 상당히 높은 빈도를 가지고 있었으므로, 다른 경우와 분리할 만한 규모라고 판단되었다. Type VII는 1,226개 경우였고, 이들 중 강우 전 날과 강우 다음 날이 무강우인 경우가 1,215개로 약 99.1%를 차지하였다. 이것은 지속시간이 1일을 넘지 않는 비교적 짧은 강우사상의 영향으로 보인다.

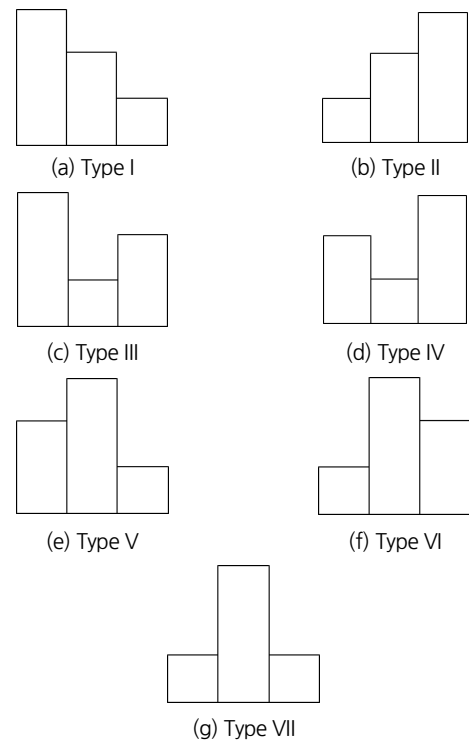


Fig. 1. Daily rainfall sequence

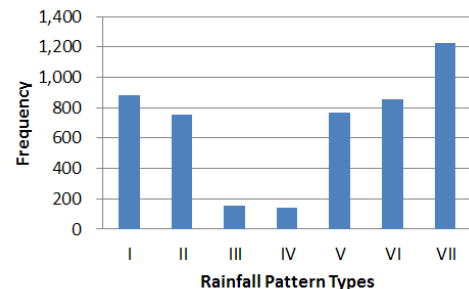


Fig. 2. Historical rainfall patterns categorized by 7 proposed sequences

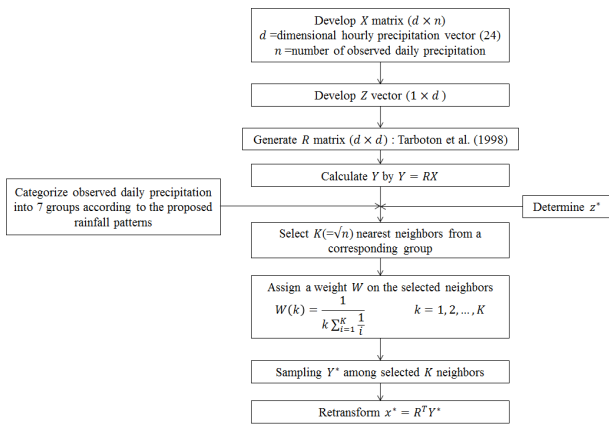


Fig. 3. Flowchart of proposed method

제안된 7개의 3일 강우패턴을 강우분해에 사용하기 위해, 현재의 일강우 자료와 하루 전과 후의 일강우 자료를 비교하여 과거자료를 Fig. 1에서 제안한 7개의 패턴으로 나누어 그룹을 만들었다. 이것은 일 강우량 값이 비슷하고, 전후일의 강우 발생 패턴도 같다면, 해당일의 시강우 분포도 유사하게 나타날 확률이 높다는 가정을 바탕으로 한다. 이 방법은 자료의 수가 적을 때는 각 그룹별 표본의 수가 부족하여 합리적인 결과가 도출되지 않을 가능성이 있으나, 표본의 수가 많다면 강우의 발달 정도에 따른 시강우 자료의 패턴을 고려할 수 있다는 장점이 있다.

전체적인 연구절차는 Fig. 3에 도시하였다.

4. 적용 및 결과

본 연구에서는 기상청 서울 관측소의 1961년에서 2012년까지 52년 간 시강우 자료를 대상으로 이를 일강우로 변환하고 변환된 일 강우자료를 다시 시강우자료로 분해하여 원본시강우자료와 재분해한 시강우자료에서 다양한 수문변량을 찾아 그 결과를 통계적으로 비교하였다.

4.1 적용자료 및 시강우 분해

먼저 시강우 자료를 일강우 자료로 변환한 결과, 전체 18,993일 중 강우일수는 4,778일로 25.1%였다. 강우 발생일 전후일의 강우량에 따라 강우 발생 패턴을 구분한 결과는 앞에서 언급한 바와 같이 Fig. 2와 같으며, 본 연구에서 새롭게 추가된 Type VII (Fig. 1(g))의 발생빈도가 25.8%로 가장 높은 것을 알 수 있다. 즉 우리나라의 강우는 2일 이상 지속되기 보다는 하루에 집중해서 내리는 경우가 많음을 의미한다. 또한

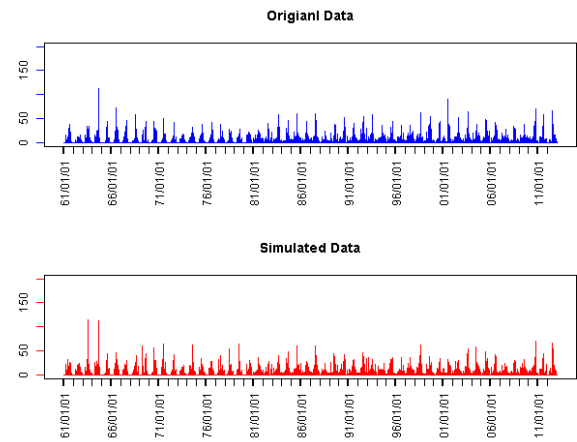


Fig. 4. Comparison of measured and disaggregated hourly rainfall

Type III과 Type IV의 발생빈도가 다른 패턴에 비해 매우 작은 것을 알 수 있다. 이 패턴들은 강우가 점차적으로 감소하다가 다시 증가하는 형태로 두 개의 강우 사상이 연결되는 경우로 보인다. 이러한 경우는 드물게 나타나는 것으로 보이지만 Type III는 156개, Type IV는 143개로 강우자료 분해를 위한 표본으로는 충분하다고 판단된다.

변환된 일자료를 앞서 제시한 방법을 통해 다시 시자료로 재분해하였다. 이때, Prairie et al. (2007)가 지정한 바와 같이 0보다 작은 값이 출현하는 경우가 발생하였다. 본 연구에 사용된 자료를 적용한 경우, 0보다 작은 값이 약 0.7% 정도 발생하였으며, 그 값은 1.1 mm 이하였다. 그러므로 Prairie et al. (2007)가 지정한 바와 같이 이것이 자료의 통계에 영향을 주기는 어렵다고 판단하였으며, 이의 보정을 위해 사전 연구에서 해당 값을 24로 나누어 일률적으로 더해주는 방법을 사용해 보았으나 24개로 나누어 각 시자료에 더해질 경우 소수점 첫째자리까지의 유효숫자로 계산한 시강우자료의 합이 절단오차(truncation error)로 인해 분해하기 전과 달라지는 경우가 발생하였다. 이에 통계적 영향을 최소화하기 위해 가장 큰 시우량에 해당 값의 절대치를 더해주는 단순한 방법을 선택하였다.

분해된 시자료와 원본자료를 Fig. 4에 나타내었다. 일자료로 변환한 수준에서는 기존의 자료와 동일한 결과를 보이는 반면, 분해된 시자료는 Fig. 4와 같이 원본자료와 전혀 다르게 나타났다.

4.2 기존 방법과 개선된 방법의 비교

원본 시자료의 통계적인 특성과 재분해된 시자료의 통계적인 특성을 살펴보기 위해 두 가지의 시자료로부터 강우사상에 대한 강우지속시간, 무강우지속시간, 강우총량 등 통계적 특성을 추출하여 서로 비교하여 보았다. 이때 두 집단이 통계

적으로 같은 확률분포에서 유래된 것인지를 판단하기 위해 비모수 검정기법인 KS 검정(Kolmogorov-Smirnov Test)을 사용하였다. KS 검정기법은 두 표본 집단이 동일한 확률분포에서 생성된 것인지를 확인할 수 있는 비모수적 검정기법이다. 귀무가설은 두 집단이 서로 같은 확률분포로부터 추출된 표본이라고 설정하였으며, 유의수준을 5%로 하여 검정을 실시하였다. 귀무가설이 모두 수락 된다면 제안된 분해 기법은 충분히 적용성이 있다고 할 수 있을 것이다.

예비분석 단계에서 두 시 자료 집단의 전체 강우자료를 이용한 경우 귀무가설이 기각되어 두 시자료 집단은 서로 다른 확률분포에서 유래한 것으로 나타났다. 하지만 1 mm 이상의 시자료에 대해서만 검정을 실시한 경우 귀무가설을 채택할 수 있어 두 자료가 같은 확률분포에서 발생한 것으로 판단할 수 있었다. 이에 1 mm를 강우와 무강우를 결정하는 기준으로

삼았다.

개선되지 않은 방법을 사용하여 분해한 경우 Table 1과 같이 전체 강우량과 강우사상에 대한 강우지속시간은 귀무가설이 채택되어 원본 시강우량 자료와 동일한 분포에서 유래한 것으로 판단할 수 있었으나, 강우사상에 대한 무강우지속시간과 강우총량은 원본 시강우자료와는 다른 것으로 나타났다. 반면에 본 연구에서 제안한 3일 강우패턴을 도입하여 개선한 방법을 적용한 경우 전체 시강우량은 물론 강우사상의 특성치들도 모두 귀무가설이 채택되어 결과의 향상이 있었음을 보여주었다. 이 검정 결과는 Table 2와 같으며, 검정과정에 나타난 각 변수별 경험적 확률분포함수의 모양은 Fig. 5와 같다. 이러한 결과는 재분해한 시자료가 원본 자료와 완벽하게 일치하는 확률분포를 갖고 있지는 않지만 특정 강우량 이상에서는 매우 유사한 확률분포를 갖고 있다는 것을 의미하며, 실

Table 1. KS Test results of disaggregated hourly rainfall without using the rainfall patterns (significance level: 5%)

Hydrological Variables	P value	Result
Hourly Rainfall Depth (over 1 mm)	0.1191	Accept
Event Duration (threshold 1 mm)	0.1696	Accept
Inter Event Time (threshold 1 mm)	0.0000	Reject
Total Rainfall of Event (threshold 1 mm)	0.0276	Reject

Table 2. KS Test results of disaggregated hourly rainfall using the proposed KNNR method with the rainfall patterns (significance level: 5%)

Hydrological Variables	P value	Result
Hourly Rainfall Depth (over 1 mm)	0.1663	Accept
Event Duration (threshold 1 mm)	0.9999	Accept
Inter Event Time (threshold 1 mm)	0.0672	Accept
Total Rainfall of Event (threshold 1 mm)	0.7887	Accept

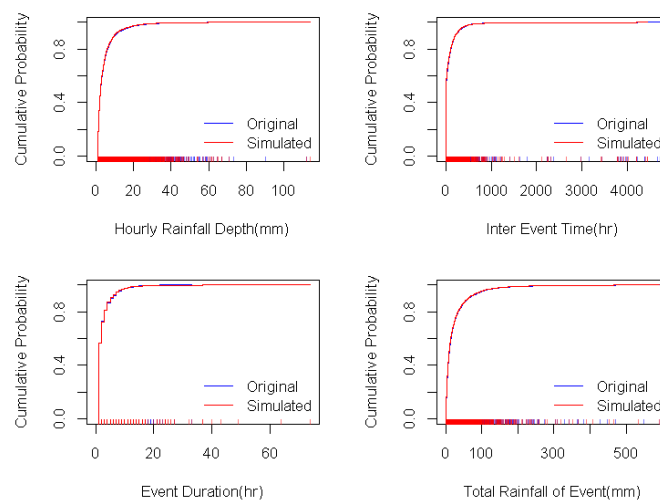


Fig. 5. Comparison of probability density functions of hydrological variables

제로 원본 시자료와 재분해된 시자료의 누가 확률분포는 Fig. 5처럼 매우 유사하다는 것을 알 수 있었다.

4.3 개선된 방법으로 분해된 결과의 적용성 평가

분해 결과의 수문학적 적용성을 살펴보기 위해 두 시자료를 이용해 지속기간별로 52개의 연최대강우강도를 구하고 그 분포를 비교 검토한 결과 Table 3와 같이 모든 경우에서 귀무가설이 채택되었다. 이것은 분해된 시자료가 빈도분석에도 충분히 활용될 수 있다는 것을 보여준다. Fig. 6은 검토에 사용된 지속기간별 연최대강우강도에 대한 경험적 확률분포 함수를 비교하여 도시한 것이다. 또한 그 결과를 이용해 원본 시자료와 생성된 시자료에 대해 I-D-F 분석을 실시해 보면 Fig. 7과 Fig. 8에서처럼 매우 유사한 결과를 보여주므로 준다.

Table 3. KS Test results of yearly maximum rainfall intensity for different durations (significance level: 5%)

Duration of Yearly Maximum Rainfall Intensity	P value	Result
1 hour	0.4173	Accept
2 hours	0.7337	Accept
6 hours	0.9698	Accept
12 hours	0.9698	Accept
24 hours	0.9979	Accept

이러한 결과를 살펴볼 때 본 논문에서 제안된 방법을 활용하여 기존 자료를 이용하여 시 강우자료의 추가적인 확보나 기

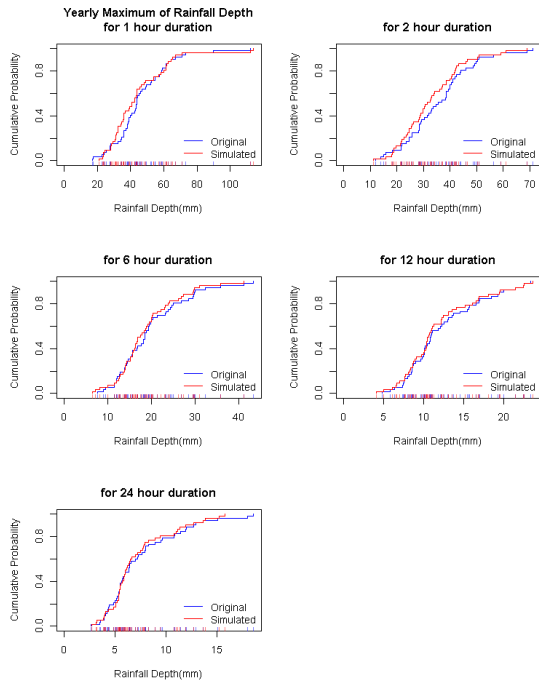


Fig. 6. Comparison of yearly maximum rainfall intensity for different durations

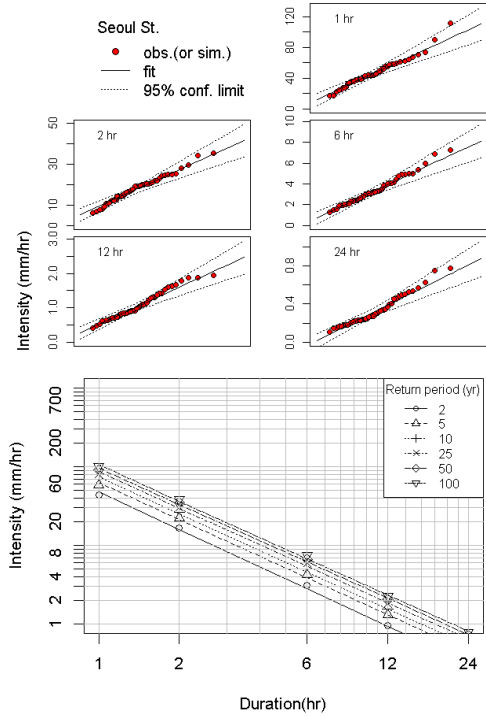


Fig. 7. I-D-F Analysis of original data

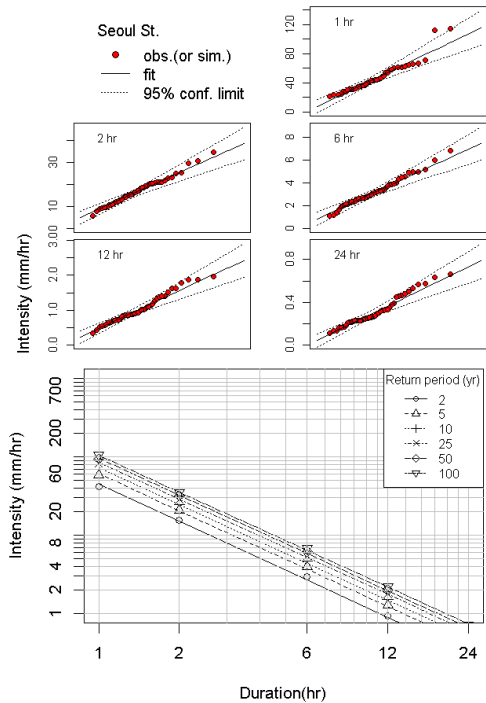


Fig. 8. I-D-F Analysis of simulated data

후변화 자료의 분해 등에 충분히 적용이 가능할 것으로 판단 된다.

5. 결론

본 연구에서는 도달시간이 짧고 첨두유출량이 큰 도시유역의 유출해석을 위해 일강우자료를 시강우자료로 분해하는 방법으로 기존의 KNNR 방법을 개선할 수 있는 방법을 제안하였다. 제안된 방법의 적용성을 평가하기 위하여 기상청 서울 관측소에서 관측된 52년간의 시강우자료를 이용하여 일강우 자료를 구하고 기존의 KNNR 방법으로 분해한 시강우자료와 제안한 KNNR 방법으로 분해한 시강우자료를 각각 원본자료와 통계학적으로 비교하여 결과가 개선되었음을 보였다.

본 연구에 적용된 KNNR 방법은 Prairie et al. (2007)에서 연유량의 월유량 분해에 적용한 수정 KNNR 방법을 Ormsbee (1989)가 제안한 강우발생 패턴을 이용하여 보완한 것이다. 본 연구에서는 Prairie et al. (2007)의 방법이 계절적 경향성이 강한 유량자료를 대상으로 개발되었다는 점을 응용하여 7개의 3일 강우발생 패턴에 따라 일강우 자료를 사전에 분리하여 해당 그룹에서만 총 일강우량이 가장 비슷한 K개의 관측값들을 선정함으로써 시강우자료에 대해서도 경향성을 부여하는 개선된 방법을 제안함으로써 통계적 경향성이 다르게 나타나는 연유량자료의 월유량자료 분해 기법을 일강우량자료의 시강우량자료의 분해에 적용할 수 있도록 개선하였다. 그 결과, 관측된 시강우자료와 그로부터 추출된 강우사상의 강우지속시간, 무강우시간, 강우총량 등에 대한 통계적 분포 특성이 원본자료와 유사하게 모의할 수 있음을 알 수 있었으며, 원본자료와 개선된 방법으로 분해된 자료를 이용하여 지속기간별 최대강우강도를 구하여 그 분포를 비교한 결과 동일한 분포에서 추출되었음을 확인하였다. 이러한 결과를 이용한다면 부족한 시강우 자료의 확장이나 기후변화 시나리오의 시간적 상세화 등에 충분히 활용할 수 있을 것으로 판단 된다.

감사의 글

본 연구는 한국연구재단 이공분야기초연구사업(NRF-2014R1A1A2056954)의 연구비 지원과 국민안전처 자연재해저감기술개발사업단(MPSS-자연-2015-79)의 연구비 지원에 의해 수행되었습니다.

References

- Kyoung, M.S., Sivakumar, B., Kim, H.S., and Kim, B.S., (2008). "Chaotic Disaggregation of Daily Rainfall from the Climate Change Scenarios" *Proceedings of Korea Water Resources Association*, pp. 178-183.
- Kim, B.S., Kim, B.K., Kyung, M.S., and Kim, H.S., (2008). "Impact Assessment of Climate Change on Extreme Rainfall and I-D-F Analysis." *Journal of Korea Water Resources Association*, Vol. 41, No. 4, pp. 379-394.
- Kim, T.J., Kwon, H.H., Lee, D.R., and Yoon, S.K. (2014). "Development of Stochastic Downscaling Method for Rainfall Data Using GCM." *Journal of Korea Water Resources Association*, Vol. 47, No. 9, pp. 825-838.
- Kumar, D.N., Lall, U., and Peterson, M.R. (2000). "Multisite disaggregation of monthly to daily streamflow." *Water Resources Research*, Vol. 36, No. 7, pp. 1823-1833.
- Lall, U., and Sharma, A. (1996). "A nearest neighbor bootstrap for resampling hydrologic time series." *Water Resources Research*, Vol. 32, No. 3, pp. 679-693.
- Lee, T., Salas, J.D., and Prairie, J. (2010). "An enhanced nonparametric streamflow disaggregation model with genetic algorithm." *Water Resources Research*, Vol. 46, W08545.
- Lee, T., and Jeong, C., (2014). "Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios." *Journal of Hydrology*, Vol. 510, pp. 182-196.
- Lee, T., Park, T., Lee, H., and Jeong, C. (2014). "Temporal Downscaling of Precipitation from Daily to Hourly Based on Nonparametric Approach: Assessment of the Climate Change Impacts on the Hourly Precipitation for the Gyeongnam Region." *Journal of Korean Society of Hazard Mitigation*, Vol. 14, No. 4, pp. 301-308.
- Liou, E.Y. (1970). OPSET-Program for computerized selection of watershed parameter values of the Stanford watershed model, Research Report No. 34, *University of Kentucky Water Resources Institute*, Lexington, KY.
- Matalas, N.C. (1967). "Mathematical assessment of synthetic hydrology." *Water Resources Research*, Vol. 3, No. 4, pp. 937-946.
- Mandelbrot, B.B., and Wallis, J.R. (1968). "Noah, Josep, and operational hydrology." *Water Resources Research*, Vol. 4, No. 5, pp. 909-918.
- Mandelbrot, B.B., and Wallis, J.R. (1969a). "Computer experiments with fractional Gaussian noises, 1. averages and variances." *Water Resources Research*, Vol. 5, No. 1, pp. 228-241.
- Mandelbrot, B.B., and Wallis, J.R. (1969b). "Computer experiments with fractional Gaussian noises, 2. Rescaled ranges and spectra." *Water Resources Research*, Vol. 5, No. 1, pp. 242-259.
- Mandelbrot, B.B., and Wallis, J.R. (1969c). "Some longrun properties of geophysical records." *Water Resources*

Research, Vol. 5, No. 2, pp. 321-340.

- Ormsbee (1989). "Rainfall disaggregation model for continuous hydrologic modeling." *Journal of Hydraulic Engineering*, Vol. 115, pp. 507-525.
- Prairie, J.R., Rajagopalan, B. Lall, U., and Fulp, T. (2007). "A stochastic nonparametric technique for space-time disaggregation of streamflows." *Water Resources Research*, Vol. 43, W03432, doi:10.1029/2005WR004721.
- Rodriguez-Iturbe, I., Mejia, J.M., and Dawdy, D.R. (1972). "Streamflow simulation, 1, A new look at markovian models, fractional Gaussian noise and crossing theory." *Water Resources Research*, Vol. 8, No. 4, pp. 921-930.
- Sharma, A., and O'Neill, R. (2002). "A nonparametric approach for representing interannual dependence in monthly streamflow." *Water Resources Research*, Vol. 138, No. 7, pp. 5-1.
- Tarboton, D.G., Sharma, A., and Lall, U. (1998). "Disaggregation procedures for stochastic hydrology based on nonparametric density estimation." *Water Resources Research*, Vol. 34, No. 1, pp. 107-119.
- Valencia, D.R., and Schaake, J.C. (1973). "Disaggregation process in stochastic hydrology." *Water Resources Research*, Vol. 9, No. 3, pp. 580-585.