



Statistical frequency analysis of snow depth using mixed distributions

Park, Kyung Woon^a · Kim, Dongwook^b · Shin, Ji Yae^c · Kim, Tae-Woong^{d*}

^aPh.D. Student, Department of Civil and Environmental System Engineering, Hanyang University, Seoul, Korea

^bM.S. Student, Department of Civil and Environmental System Engineering, Hanyang University, Seoul, Korea

^cPostdoctoral Fellow, Department of Civil and Environmental Engineering, Hanyang University, Ansan, Korea

^dProfessor, Department of Civil and Environmental Engineering, Hanyang University, Ansan, Korea

Paper number: 19-076

Received: 2 October 2019; Revised: 21 November 2019; Accepted: 21 November 2019

Abstract

Due to recent increasing heavy snow in Korea, the damage caused by heavy snow is also increasing. In Korea, there are many efforts including establishing disaster prevention measures to reduce the damage throughout the country, but it is difficult to establish the design criteria due to the characteristics of heavy snow. In this study, snowfall frequency analysis was performed to estimate design snow depths using observed snow depth data at Jinju, Changwon and Hapcheon stations. The conventional frequency analysis is sometime limited to apply to the snow depth data containing zero values which produce unrealistic estimates of distribution parameters. To overcome this problem, this study employed mixed distributions based on Lognormal, Generalized Pareto (GP), Generalized Extreme Value (GEV), Gamma, Gumbel and Weibull distribution. The results show that the mixed distributions produced smaller design snow depths than single distributions, which indicated that the mixed distributions are applicable and practical to estimate design snow depths.

Keywords: Frequency analysis, Mixed distribution, Maximum snowfall depth, Return period

혼합분포함수를 적용한 최심신적설량에 대한 수문통계학적 빈도분석

박경운^a · 김동욱^b · 신지예^c · 김태웅^{d*}

^a한양대학교 대학원 건설환경시스템공학과 박사과정, ^b한양대학교 대학원 건설환경시스템공학과 석사과정,

^c한양대학교(ERICA) 건설환경공학과 박사후연구원, ^d한양대학교(ERICA) 건설환경공학과 교수

요 지

최근 우리나라에서 폭설이 증가하고 있으며, 이로 인한 피해액 또한 증가하고 있다. 우리나라는 전국적으로 폭설로 인한 피해를 줄이기 위해 내설 설계기준 마련 등의 노력을 하고 있으나, 강설 자료의 특성으로 기준 설정에 어려움이 있다. 본 연구에서는 우리나라 남부 지역에 있는 진주, 창원, 합천 지점의 적설량에 대한 수문통계학적 빈도분석을 수행하여 최심신적설량에 대한 설계수문량을 경량적으로 산정하였다. 자료의 특성상 연도별 측정값이 '0'인 경우가 존재하여 기존의 빈도분석 방법을 적용할 경우 매개변수의 추정이 불가능한 상황도 발생한다. 이러한 문제를 해결하기 위하여 혼합분포함수를 이용하였고, 분포모형으로는 대수정규, 일반화 파레토, 일반화 극치, 감마, 겐벨, 와이블 분포를 적용하였다. 적용 결과, 단일분포함수를 적용할 때 보다 혼합분포함수를 적용할 때 확률적설심이 더 작게 산정되었으며, 전체적으로 관측값이 간헐적으로 나타나는 지점에서 혼합분포함수의 적용성이 우수한 것으로 판단된다.

핵심용어: 빈도분석, 혼합분포함수, 최심신적설, 재현기간

*Corresponding Author. Tel: +82-31-400-5184
E-mail: twkim72@hanyang.ac.kr (T.-W. Kim)

1. 서론

최근 기후변화의 영향으로 우리나라 전 지역에 걸쳐 자연재해로 인한 피해가 증가하고 있다. 이러한 현상의 주된 원인인 지구온난화는 범지구적 기온 상승을 뜻하며, 그 결과 여름철과 겨울철에 각각 홍수와 폭설로 나타나 지구 곳곳에 큰 피해를 입히고 있다(Korean Society of Hazard Mitigation, 2012). 지구온난화는 북극지역 얼음의 해빙(解氷)을 유발하며, 이는 구름의 대량 생성과 북반구 지역의 강설로 이어진다. 북반구의 눈과 구름은 햇빛을 반사시켜 지표면의 온도를 떨어뜨리고, 이 때문에 늘어난 찬 공기는 시베리아 고기압을 강하게 발달시킨다. 이렇게 발생한 시베리아 고기압은 한반도로 내려오면서 폭설을 유발한다(Han, 2014). 우리나라에서는 거의 매해 폭설로 인한 피해가 발생한다. 특히, 2001년 1월 폭풍을 동반한 폭설의 영향으로 전국적으로 약 6,600억원의 피해액이 발생했고(Yu *et al.*, 2016), 2011년 2월 삼척시에는 110 cm 가 넘는 폭설이 내려 건물 지붕이 붕괴되고 도시가 마비되는 피해가 있었다(Kwon *et al.*, 2016).

대설은 짧은 시간에 많은 양의 눈이 오는 기상현상을 의미한다. 국민안전처에 따르면, 대설에 의한 피해는 눈이 쌓여서 발생하는 적설 피해, 눈 압력에 의해 발생하는 설압 피해, 가파른 경사면의 적설이 미끄러져 발생하는 눈사태 피해, 젖은 눈이 송전선 등의 가설물에 부착되어 발생하는 착설 피해, 도로에 얼은 눈으로 인한 교통사고 피해, 그리고 대설 발생 후 관리 소홀로 인한 피해 등 크게 6가지로 나뉜다(Ministry of Public Safety and Security, 2015). 우리나라의 경우, 2008년부터 2017년까지의 원인별·연도별 총 피해액 중 태풍과 호우로 인한 피해액은 약 3조 1천억원으로 전체의 88% 수준이고, 대설로 인한 피해액은 약 2천 2백억원으로 전체의 6% 정도의 수준이다(Ministry of the Interior and Safety, 2018). 그러나 최근 증가하고 있는 대설로 인한 피해로, 2017년 행정안전부는 강설일수, 최심신적설 및 신적설 자료를 활용하여 대설 재해 및 대설 위험도에 대한 분석을 수행할 것을 “시·군·동 풍수해저감종합계획 세부수립기준”에서 제시하였다. 특히, 하루 동안 쌓인 눈의 깊이를 의미하는 최심신적설은 비닐하우스의 설계와 제설 작업의 기준이 되기 때문에(Ministry of Public Safety and Security, 2015), 이에 대한 정량적인 분석이 필요하다.

우리나라에서는 건축물에 대하여 내설설계기준이 설정되어 있다. 2016년 고시된 “건축구조기준”과 2014년 고시된 “원예·특작시설 내재해형 기준”에서는 각각 100년, 30년 빈도의 적설심을 기준으로 설계할 것을 제시하고 있다. 하지만, 해당 빈도에 해당하는 적설심을 산정하는 기준 및 가이드라인은 제시되어 있지 않은 실정이다. 이 때문에 대부분의 연구에

서는 수문자료를 활용하는 빈도분석 기법과 동일한 방법으로 대설의 발생빈도에 대한 정량적인 분석을 수행하고 있다. Lee *et al.* (2007)은 기상청 52개 관측지점의 최심신적설량 자료를 이용한 최적분포형 결정 및 재현기간별 확률적설량을 산정하였으며, Gamma 분포 모형이 대체적으로 적용성이 높다는 것을 확인한 바 있다. Kim *et al.* (2010)은 기상청 61개 관측지점의 최심신적설량 자료를 이용해 최근 적설량의 확률빈도규모를 고려한 확률적설량도를 작성하였으며, 남부지역에서 북부지역으로, 서부지역에서 동부지역으로 갈수록 확률적설량이 높아진다는 것을 확인하였다. 또한, Yu *et al.* (2016)은 적설심 빈도분석을 수행하여 온실의 적설심 기준을 평가하였으며, 30년 빈도 확률적설심과 설계적설심을 비교한 결과 230개 행정구역 중 77개 행정구역의 설계적설심이 적정한 것을 확인하였다. 대부분의 연구 사례에서는 최심신적설을 활용하여 빈도분석을 수행하고 있으며, Lee *et al.* (2007), Yu *et al.* (2016)과 같이 단일분포함수를 적용한 빈도분석 방법을 수행한 경우가 대부분이다. 하지만 우리나라 남부지방 중 일부 지역에서는 연중 눈이 내리지 않는 경우가 있어 연속적으로 존재하는 강우 자료와는 다르게 ‘0’ 값이 존재하거나, 자료가 결측되는 경우가 발생한다. 일반적으로 특정년도의 값이 결측된 경우에는 해당 자료를 제외하고 빈도해석을 수행한다. 하지만, 관측된 자료에 ‘0’ 값이 존재할 경우 자료의 통계적인 특성들은 매우 민감하게 반응하며, ‘0’ 값을 제외하고 추정된 통계특성과 크게 다를 수 있다(Kim and Yoo, 2007; Park and Chung, 2019).

Haan (2002)은 0 값이 포함되어 있는 수문자료의 통계적 분석을 수행하는 방법으로 세가지를 제시한 바 있다. 첫 번째로는 ‘0’ 값을 포함한 모든 관측 자료에 일정하게 작은 값을 더해 주는 것으로, 적합 결과의 신뢰성이 낮다는 문제점이 있다(Wang and Singh, 1995). 두 번째 방법은 ‘0’인 자료를 제외한 자료만을 활용하여 분석을 수행하는 것으로, 이는 ‘0’ 값이 가지고 있는 통계적 특성을 무시하는 한계가 있다. 마지막으로, Jennings and Benson (1969)이 제시한 전확률 이론을 기반으로 ‘0’인 값들의 사상과 ‘0’이 아닌 값들의 사상을 각각의 독립된 모집단으로 판단하여 분석하는 방법이다. 즉, 두 모집단들의 발생 비율을 고려하여 각각의 적합한 확률분포모형을 적용하는 방법인 혼합분포를 활용하는 방법으로, ‘0’ 값을 포함한 자료로 홍수빈도해석을 수행할 경우, 낮은값에서는 첫 번째 방법보다 적합성이 높다는 것을 확인하였다. 그 밖에도 Woo and Wu (1989), Wang and Singh (1995), Strupczewski *et al.* (2003), Wei and Song (2019)가 전확률 이론을 기반으로 하는 혼합분포 모형을 활용하여 ‘0’ 값을 포함한 수문자료의 통계적 분석을 수행하였다. 해당 연구내용을 바탕으로, 본 연구에서는 ‘0’ 값이 포함되어 있는 관측 최심신적설량의 수문통계

학적 빈도분석을 위해 혼합분포모형을 활용하는 방법을 제안하였다.

혼합분포함수는 여러 원인에 의해 일어나는 현상을 분석하는데 적합하며(Haan, 2002), 적은 양의 자료를 정량적으로 비교 분석할 수 있는 장점이 있다(Shin *et al.*, 2013). 예를 들어, 극치 풍속자료를 활용한 빈도해석에서는 풍속자료의 높은 변동성과 지역성을 반영하기 위하여 기존의 확률분포형이 아닌 혼합분포모형을 활용한 빈도해석을 수행하였다(Shin *et al.*, 2018). 이 밖에도 Shin *et al.* (2013)은 혼합분포함수를 적용하여 드물게 발생하는 지진자료에 대한 빈도분석을 수행하여 지진의 발생빈도를 추정할 바 있다. 수자원 분야에서는 Yoo and Kim (2004)이 혼합분포함수를 적용하여 빈도분석을 실시하였을 때 강수기간과 무강수기간을 동시에 고려하여 강우의 간헐성을 표현하는데 유리하다는 것을 예증하였으며, Shin and Lee (2014)는 수문자료에 메타-휴리스틱 최우도법으로 혼합분포모형의 빈도해석을 수행하는 방법을 제안하였다. Choi *et al.* (2018)은 혼합 Gumbel 분포 및 Bayesian 매개변수 추정 기법을 기반으로 강우빈도해석을 수행하였다. 이처럼 최근 수문분야에서도 혼합분포함수를 활용함으로써 자료의 통계적 특성을 효과적으로 반영하고 있다. 이에 따라 본 연구에서는 우리나라의 남부지역에서 눈이 내리지 않았던 해의 자료를 반영하기 위하여 혼합분포함수를 이용한 최심신적설량 자료의 수문통계학적 빈도분석을 수행하였다.

2. 연구방법

2.1 혼합분포함수

혼합분포함수란 두 가지 이상의 확률분포함수를 결합한 것으로, 일반적으로 이산분포와 연속분포의 혼합으로 이루어진다. 혼합분포함수는 여러 개의 뚜렷한 원인에 의한 특정 사상을 분석하는데 효과적이다(Haan, 2002). Kedem *et al.* (1990)은 두 가지 확률분포함수를 결합한 혼합분포함수를 Eq. (1)과 (2)와 같이 제시하였다.

$$G(x) = pF(x) + (1-p)H(x) \quad (1)$$

$$p = \frac{n}{N} \quad (2)$$

여기서, p 는 연속분포와 관련된 사상의 발생 확률, n 은 관측된 '0' 값이 아닌 자료 수, N 은 자료의 전체 관측기간을 나타내며, $F(x)$ 는 특정조건 하의 연속분포에 대한 누가확률밀도함수,

$H(x)$ 는 불연속적인 부분을 표현하는 이산분포에 대한 누가확률밀도함수를 나타낸다. $H(x)$ 는 Eq. (3)과 같은 계단함수(step function)로 정의된다(Yoo *et al.*, 2005).

$$H(x) = \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{if } x > 0 \end{cases} \quad (3)$$

본 연구의 경우, 전체 관측 자료는 적설이 존재하는 기간의 자료, 적설이 존재하지 않는 기간의 자료, 이렇게 두 가지 모집단으로 이루어져 있다. 적설 자료가 존재하는 기간은 $H(x)$ 가 0, 존재하지 않는 기간에는 $H(x)$ 가 1로 두 확률분포 모형을 결합한 $G(x)$ 를 Eq. (4)와 같이 표현할 수 있다.

$$G(x) = \begin{cases} pF(x), & \text{if } x = 0 \\ pF(x) + (1-p), & \text{if } x = 1 \end{cases} \quad (4)$$

여기서 $F(x)$ 는 본 연구에서 적용한 Gamma, Generalized Pareto (GP), Generalized Extreme Value (GEV), Gumbel, Lognormal, Weibull 확률분포함수의 누가분포함수이다.

2.2 적설 자료를 활용한 통계학적 빈도분석

빈도분석이란 확률통계학적 원리를 이용하여 확률변수 크기와 발생확률의 관계를 분석하는 것이며, 수문자료를 확률론적 빈도분석에 적용하는 것을 수문학적 빈도분석이라 한다. 수문자료의 빈도분석은 구축한 자료의 독립성, 정상성, 동질성을 확인한 후 확률분포의 특성을 나타내는 매개변수를 추정하고, 매개변수를 통해 구한 확률분포모형이 자료에 적합한지 판단하기 위해 적합도 검정을 수행한 후 재현기간에 따른 사상의 빈도를 나타내는 절차로 진행된다.

Ministry of Public Safety and Security (2015)는 대설위험도 분석 시에 다양한 설해 관련 자료(최심적설, 최심신적설) 중에서 최심신적설을 활용하여 빈도분석을 수행할 것을 권장하였다. 따라서, 본 연구에서는 기상청에서 제공하는 최심신적설(하루에 가장 많은 눈이 쌓였을 때의 깊이) 자료를 활용하였다. 빈도분석을 수행하는 연구 대상 지점(Fig. 1)은 눈이 많이 내리지 않는 우리나라 남부 지방의 진주(Jinju), 창원(Changwon), 합천(Hapcheon) 지점으로 자료의 연한은 각각 1987~2017(44년), 1971~2014(31년), 1971~2008(38년)이다. 창원 및 합천의 경우 각각 2015, 2009년부터 2017년도까지는 자료가 없어, 본 연구에서는 '0' 값이 아닌 관측값의 미계측으로 판단하였다. 지점 별 자료기간 내에서도 Fig. 2와 같이 자료가 없는 경우가 있으며, 해당 기간의 경우에는 관측값이 '0'인 것으로 판단하였다.

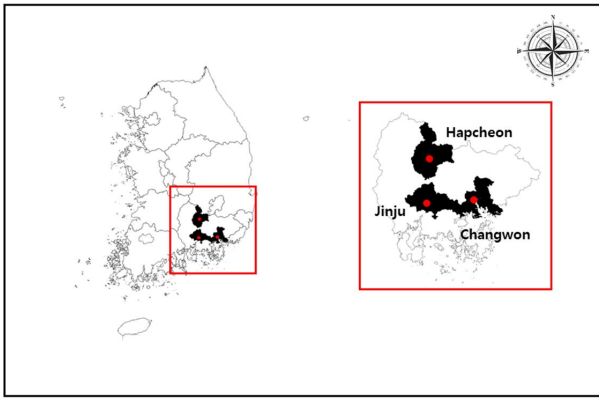


Fig. 1. Weather stations used in this study

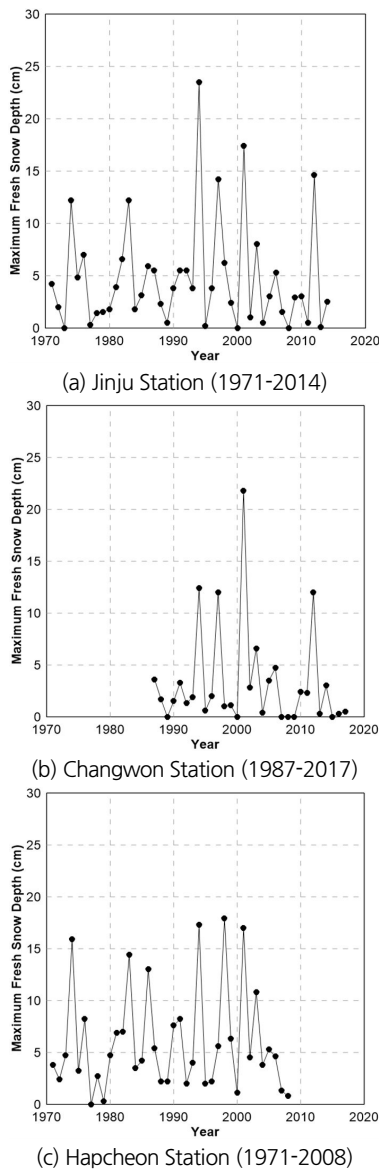


Fig. 2. Maximum fresh snow depth of Jinju, Changwon, Hapcheon stations

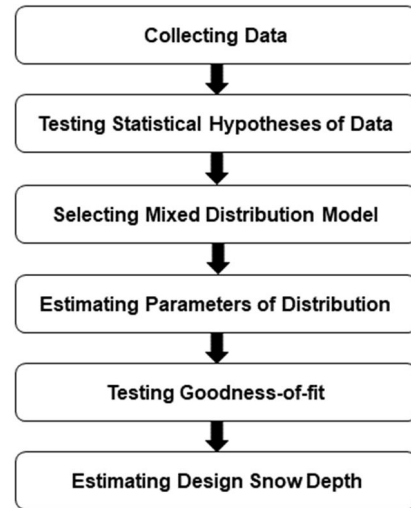


Fig. 3. Process of frequency analysis for snow depth

일반적으로 수행되는 홍수(강우)빈도분석의 경우 대부분의 자료에는 ‘0’값이 포함되지 않기 때문에, 구축한 자료의 독립성, 정상성, 동질성을 확인한 후, 확률분포모형의 매개변수를 추정하고 분포모형의 적합성을 검토한 다음, 재현기간에 따른 사상의 빈도를 산정하는 절차로 진행된다. 하지만, 자료에 ‘0’값이 포함될 경우 다음과 같은 분석의 어려움이 존재한다. 자료에 ‘0’값이 존재하면 로그함수의 값을 산정하지 못하기 때문에 일반적인 빈도분석에서 사용되는 확률밀도함수를 적용하지 못할 수도 있다. 또한, 일부 분포모형의 경우 ‘0’값이 포함될 경우, 최우도법으로는 매개변수 추정이 불가능한 경우가 발생하게 된다. 이 때문에 본 연구에서는 혼합분포함수를 적용하여 ‘0’값에 대한 모집단을 분류하여 빈도분석을 수행하는 방법을 제시하며, 적용방법은 Fig. 3에 나타냈다. Fig. 3의 혼합분포를 활용한 빈도분석 방법은 일반적인 빈도분석 과정과 비슷하게 진행된다. 두 방법의 차이점은 기존의 빈도분석 방법은 연최대치시계열 자료를 그대로 빈도분석에 적용했던 것과 다르게 전체 자료 중에서 ‘0’인 자료를 제외하는 과정이 추가되어 나머지 자료만을 활용하여 빈도분석이 수행된다는 것이다.

3. 적용 및 연구 결과

3.1 기초통계분석

적설량 자료를 활용한 빈도분석에 앞서 자료의 통계적 특성에 대한 분석을 수행하여 그 결과를 Table 1에 나타냈다. Table 1과 같이 창원 지점은 다른 두 지점과 비교했을 때 자료의 연한이 짧고 ‘0’값이 많기 때문에 p 값이 작게 산정되었다.

Table 1. Basic statistics of maximum fresh snow depth in Jinju, Changwon, Hapcheon stations

	Jinju	Changwon	Hapcheon
p	0.932	0.806	0.974
Mean	4.698	3.313	6.023
Standard Deviation	5.106	4.906	4.988
Skewness	1.808	2.267	1.133
Kurtosis	3.151	4.845	0.215

3.2 혼합분포함수 기반의 빈도분석 결과

3.2.1 매개변수 추정 결과

본 연구에서는 최우추정법을 적용하여 확률분포의 매개변수를 추정하였다. 최우추정법은 주어진 표본자료가 나올 확률이 최대가 되도록 확률분포형의 매개변수를 추정하는 방법으로 n 개의 표본자료 확률밀도함수 $f(x_n; \theta)$ 를 가진 최우추정치 θ 와 우도함수 $L(\theta)$ 의 관계는 Eq. (5)로 표현된다.

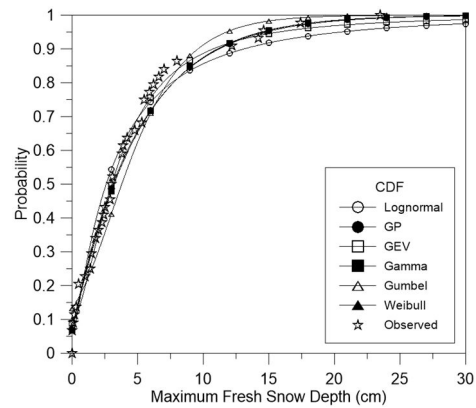
$$L(\theta) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta) \quad (5)$$

본 연구에서는 Eq. (1)에서 적설량이 존재하는 기간의 확률분포함수인 $F(x)$ 에 Gamma, Generalized Pareto (GP), GEV, Gumbel, Lognormal, Weibull 분포를 적용하였다. 확률분포함수의 매개변수를 추정한 결과, 진주, 창원, 합천 지점의 대설자료에 대한 혼합분포함수 $G(x)$ 의 누적분포함수는 Fig. 4와 같이 도시된다.

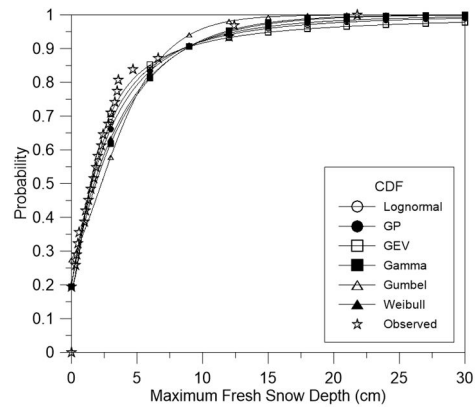
3.2.2 혼합분포함수 적합도 검정

본 연구에서는 혼합분포함수의 적합성 검토를 위해 K-S 검정과 χ^2 검정, 그리고 Cramer Von Mises (CVM) 검정을 적용하였다. K-S 검정은 관측된 자료의 누적확률 $F_n(x)$ 와 이론적인 누적확률 $F(x)$ 의 차이 D_n 를 구하고, 이 차이를 한계치와 비교하는 방법이다. χ^2 검정은 관측자료의 주상도와 이론확률분포의 차이를 유의수준에 따른 한계값과 비교하는 방법이다. CVM 검정은 K-S 검정과 비슷하게 누적확률을 이용하며, 관측수문자료 x_n 이 누적확률 $F_x(x; \theta)$ 의 확률분포형을 모집단으로 한다는 것을 검정하는 방법이다. K-S 검정은 Eq. (6), χ^2 검정은 Eq. (7), CVM 검정은 Eq. (8)과 같이 표현될 수 있다. χ^2 검정에서 χ^2 는 검정통계량, n 은 관측값의 개수, k 는 계급구간의 개수, p_i 는 상대도수, $n p_i$ 는 절대도수를 뜻하며, CVM 검정에서 W 는 검정통계량, n 은 관측값의 개수, θ 는 관측수문자료에서 도출된 매개변수를 나타낸다.

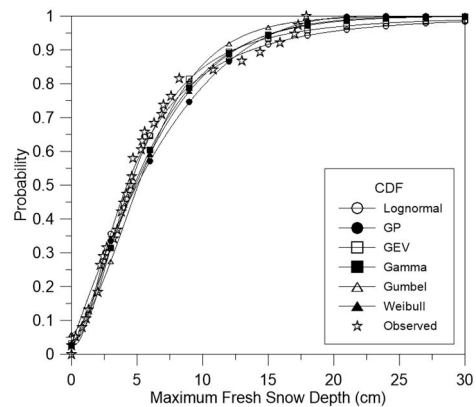
$$D_n = \max |F_n(x) - F(x)| \quad (6)$$



(a) Jinju Station



(b) Changwon Station



(c) Hapcheon Station

Fig. 4. Cumulative distribution functions of mixed distributions

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} \quad (7)$$

$$W = \frac{1}{12n} + \sum_{i=1}^n [F_x(x; \theta) - \frac{2i-1}{2n}]^2 \quad (8)$$

본 연구에서 제시한 혼합분포함수를 적용한 빈도분석 결과의 적절성 검토를 위해 단일분포함수를 적용한 결과와 비교를 수행하였다. 단일분포함수를 적용할 경우, 2변량 Lognormal 분포함수는 자료에 '0'이 존재할 경우 활용이 불가능하며 (Chow *et al.*, 1988), GP 및 Weibull 분포함수는 '0'값의 영향으로 최우추정법을 이용할 때 매개변수의 추정이 되지 않는다. 이 때문에 전체 6개의 함수 중에서 3개의 단일분포함수(GEV, Gamma, Gumbel)만이 혼합분포함수와 비교가 가능하였다.

유의수준 5%에 대한 검정 결과, 진주와 합천 지점에 대한 Gumbel 분포함수의 K-S 검정을 제외한 나머지 검정에 대해서 적합성을 확인하였다. 혼합분포함수는 본 연구에서 활용한 대부분의 분포함수에 대해 검정을 통과하였고, 창원 지점에 대한 Gumbel 분포함수의 K-S 검정과 Gamma 분포함수의 CVM 검정 결과만 기각되었다. 3가지 검정 기법 중에서 Lee *et al.* (2007)을 바탕으로 χ^2 검정 기법에 대하여 가장 적합성이 높은 분포모형은 진주 지점은 혼합 Gumbel (p-value 0.798), 창원 지점 Gamma (p-value 0.875), 합천 지점 혼합 GEV (p-value 0.954)였다. 하지만, 본 연구는 혼합분포모형의 적절 자료에 대한 빈도해석의 적용 가능성을 검토하는 것을 목적으로 하므로 최적의 분포 모형을 제시하기 보다는 전체 분포모형을 활용하여 빈도곡선을 작성하고 그 결과를 비교하였다.

3.2.3 재현기간에 따른 빈도곡선

본 연구에서는 Eqs. (9)~(11)과 같이 최심신적설 자료에 대한 빈도분석을 통해 재현기간에 따른 확률적설심을 산정하였다.

$$\frac{1}{T} = 1 - G(x; \mu_1, \sigma_1, p) = p - pF(x; \mu_1, \sigma_1) \quad (9)$$

$$F(x; \mu_1, \sigma_1) = 1 - \frac{1}{pT} \quad (10)$$

$$x_T = F^{-1}\left(1 - \frac{1}{pT}\right) \quad (11)$$

여기서 T 는 재현기간, $G(x)$ 는 비초과확률이다.

기존 방법의 단일분포함수(GEV, Gamma, Gumbel)와 혼합분포함수(Mixed Lognormal, Mixed GP, Mixed GEV, Mixed Gamma, Mixed Gumbel, Mixed Weibull)의 빈도분석 결과인 재현기간에 따른 확률적설심을 Table 2에 나타냈다. 진주와 창원 지점의 혼합 GEV, Gamma 분포함수가 단일 GEV, Gamma 분포함수와 비교했을 때 더 작은 값의 확률적설심을 가지며, 합천 지점의 경우 혼합 Gamma, Gumbel 분포함수가 단일 Gamma, Gumbel 분포함수보다 더 작은 값의 확률적설심을 갖는다.

3.3 단일분포함수를 적용한 빈도분석 결과와의 비교

앞에서 설명한 바와 같이 '0'값이 포함된 자료의 경우, Lognormal, GP, Weibull 분포함수는 적용하기 어렵기 때문에, Gamma, GEV, Gumbel 분포의 매개변수만을 추정해 적합도 검정을 수행할 수밖에 없었다. 하지만, 혼합분포함수를 활용할 경우 전체 자료에서 '0'인 값의 모집단과 '0'이 아닌 값의 모집단으로 구분하여 분석이 수행되므로 전체 6개의 분포모형의 적용이 가능하다. 이 때문에 혼합분포함수의 매개변수 추정 및 적합성 검정이 가능하고, 기존의 방법보다 더 많은 함수들의 적용이 가능하다. 이 중에서 단일 및 혼합분포함수 모두 적용된 3가지(Gamma, GEV, Gumbel) 분포함수의 빈도분석 결과를 내설설계기준과 비교하였다.

현재 우리나라의 내설 설계 기준으로는 건축구조기준 그리고 원예 및 특작시설 내재해형 기준 두 가지가 존재하며, 건축물 설계는 100년 빈도의 적설심과 적설심에 따른 평균 단위중량을 곱한 기본지상적설하중을 기준으로 하며, 내재해형 온실 설계는 30년 빈도의 적설심을 기준으로 한다 (Chung, 2017). 건축물 설계 기준인 100년 빈도의 적설심은 제시된 값이 없어 본 연구에서는 원예 및 특작 시설 내재해형 온실 설계 기준인 재현기간 30년의 적설심과 비교하여 Table 3에 제시하였다. 진주와 창원의 단일, 혼합 GEV 분포함수를 제외하고는 대부분의 경우에서 온실 설계 기준보다 적설심이 낮게 산정된다는 것을 확인할 수 있다. Yu *et al.* (2016)이 수행한 단일분포모형 기반의 적설빈도분석에서도 30년 빈도의 확률적설심이 진주, 창원, 합천 지점 각각 18 cm, 19 cm, 19 cm로 온실 설계 기준인 20 cm, 20 cm, 22 cm보다 더 작게 산정된 것을 확인할 수 있다. GEV 분포는 극치값에 최적화된 분포모형으로, 추정된 매개변수를 활용한 분포모형을 적용할 경우 비현실적인 값 즉 과다 추정되는 문제가 발생할 수 있다(De Paola *et al.*, 2018). 창원 지점의 경우 전체 31개 값 중에서 '0'값이 6개로 다른 지점들과 비교하였을 때 '0'값이 가장 많이 포함되어 있어 이에 대한 영향으로 단일 GEV 분포함수를 활용할 경우 확률적설심이 과다 산정된 것으로 판단된다. 창원 지점의 경우 GEV 혼합분포함수를 활용함으로써 온실 설계 기준 적설심에 가까운 값을 제시할 수 있었다.

4. 결론

본 연구에서는 진주, 창원, 합천 지점 적설자료의 분석을 위해 수문통계학적 빈도분석을 수행하였다. 해당 지역의

Table 2. Comparison of design snow depths for single and mixed distributions

Station	Distribution		Return Period (year)						
			10	20	30	50	80	100	200
Jinju	Single	GEV	11.1	16.8	21.1	27.9	35.7	40.1	57.1
		Gamma	11.2	14.9	17.1	19.8	22.4	23.6	27.3
		Gumbel	9.5	11.6	12.9	14.5	15.9	16.6	18.7
	Mixed	Lognormal	13.1	20.6	25.9	34.0	42.8	47.4	64.3
		GP	11.2	14.8	16.8	19.5	21.9	23.0	26.6
		GEV	10.8	15.8	19.5	25.1	31.5	35.0	48.2
		Gamma	11.1	14.5	16.4	18.9	21.2	22.2	25.6
		Gumbel	9.6	11.8	13.1	14.6	16.1	16.8	18.9
Changwon	Single	GEV	13.9	36.0	61.9	122.0	227.0	304.5	756.6
		Gamma	9.1	13.2	15.6	18.7	21.7	23.1	27.6
		Gumbel	7.0	8.8	9.8	11.0	12.2	12.7	14.4
	Mixed	Lognormal	8.5	13.3	16.8	21.9	27.5	30.4	41.1
		GP	8.6	12.6	15.4	19.3	23.4	25.5	32.9
		GEV	8.6	15.5	21.7	32.9	48.1	57.5	99.8
		Gamma	8.7	11.7	13.5	15.7	17.8	18.8	21.8
		Gumbel	7.6	9.5	10.6	12.0	13.2	13.8	15.7
		Weibull	8.8	12.1	14.0	16.5	18.8	20.0	23.4
		Hapcheon	Single	GEV	12.4	16.6	19.5	23.6	27.9
Gamma	12.6			15.8	17.7	20.0	22.1	23.1	26.2
Gumbel	11.3			13.7	15.1	16.8	18.4	19.1	21.4
Mixed	Lognormal		13.7	19.1	22.7	27.8	32.9	35.6	44.6
	GP		13.2	15.5	16.6	17.8	18.8	19.1	20.2
	GEV		12.4	16.8	19.9	24.4	29.1	31.6	40.6
	Gamma		12.4	15.5	17.2	19.4	21.3	22.3	25.2
	Gumbel		11.3	13.6	15.0	16.7	18.2	19.0	21.2
	Weibull		12.5	15.4	16.9	18.9	20.6	21.4	23.9

Table 3. Comparison of 30-Year snow depth and design snow depth

Station	Distribution	30-year snow depth in the design guideline (A) (cm)	30-year snow depth estimated in this study (B) (cm)	Difference (A-B) (cm)
Jinju	GEV	20	21.1	-1.1
	Gamma	20	17.1	+2.9
	Gumbel	20	12.9	+7.1
	Mixed GEV	20	19.5	+0.5
	Mixed Gamma	20	16.4	+3.6
	Mixed Gumbel	20	13.1	+6.9
Changwon	GEV	20	61.9	-41.9
	Gamma	20	15.6	+4.4
	Gumbel	20	9.8	+10.2
	Mixed GEV	20	21.7	-1.7
	Mixed Gamma	20	13.5	+6.5
	Mixed Gumbel	20	10.6	+9.4
Hapcheon	GEV	22	19.5	+2.5
	Gamma	22	17.7	+4.3
	Gumbel	22	15.1	+6.9
	Mixed GEV	22	19.9	+2.1
	Mixed Gamma	22	17.2	+4.8
	Mixed Gumbel	22	15.0	+7.0

경우, 일부 기간의 최심신적설량 관측값이 '0'값으로 기록되어 Lognormal 분포함수를 적용한 빈도분석이 불가능하고, Weibull, GP 분포함수의 최우도법을 활용한 매개변수 추정이 불가능하다. 또한, '0'인 관측값의 수가 많을 경우, 자료의 통계적 특성이 달라지므로 빈도분석시 과소 혹은 과다 추정되는 문제가 발생할 수 있어, 본 연구에서는 혼합분포함수를 활용한 빈도분석을 제안하였다. 기존의 빈도분석 방법으로 도출된 결과와 혼합분포함수를 적용한 방법으로 도출된 결과를 통해 재현기간별 확률적설심을 비교하였으며, 단일분포함수를 적용할 경우보다 혼합분포함수를 적용할 경우가 확률적설심이 낮게 산정되는 경향이 있다는 것을 확인하였다. 이는 '0'인 자료의 사상을 분리하여 확률분포모형을 적용하였기에 나타나는 결과로 판단된다. '0'값의 수가 1개(전체 38개 값, '0'값은 전체의 2.6%)로 가장 적은 합천 지점의 경우 단일분포함수와 혼합분포함수의 결과가 큰 차이는 없었다. 하지만 '0' 값의 수가 6개(전체 31개 값, '0'값은 전체의 19.4%)로 가장 많은 창원 지점의 경우 '0'의 영향으로 일부 분포모형 중에서 단일분포함수로 빈도분석을 적용한 결과가 과다 산정되는 경향이 나타났다. 때문에 혼합 GEV 분포함수를 활용함으로써 과다 산정되는 추정값을 낮출 수 있었다. 본 연구에서 수행된 분석의 경우는 일부 지점으로 한정되어 있으나, 관측값이 보다 간헐적으로 나타나는 지점에서는 혼합분포함수의 적용성이 우수할 것으로 판단된다.

현재 내설 설계 기준에서는 빈도해석의 절차가 수립되어 있지 않고 연구자마다 활용하는 자료도 달라 도출되는 빈도별 적설심 또한 매우 상이하다(Chung, 2017). 따라서 본 연구의 결과는 '0'값이 존재하는 대설 자료의 분석 방법 중의 하나으로써 내설 설계 기준의 수립 및 개선에 활용 가능할 것으로 판단된다.

감사의 글

본 연구는 정부(행정안전부)의 재원으로 재난안전기술개발사업단의 지원을 받아 수행된 연구임. [MOIS-재난-2015-05]

References

- Choi, H.G., Uranchimeg, S., Kim, Y.T., and Kwon, H.H. (2018). "A bayesian approach to gumbel mixture distribution for the estimation of parameter and its use to the rainfall frequency analysis." *Journal of The Korean Society of Civil Engineers*, KSCE, Vol. 38, No. 2, pp. 249-259 (in Korean).
- Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied hydrology*. McGraw-Hill, USA.
- Chung, G.H. (2017). "Design criteria for reducing snowfall damages." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 17, No. 1, pp. 35-43 (in Korean).
- De Paola, F., Giugni, M., Pugliese, F., Annis, A., and Nardi, F. (2018). "GEV parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in african test cities." *Hydrology*, USGS, Vol. 5, No. 28, pp. 1-23.
- Haan, C.T. (2002). *Statistical methods in hydrology*. Iowa State University Press, Ames, Iowa, USA.
- Han, W.S. (2014). *The increase of heavy snow vulnerable area due to climate change and its countermeasures*. Korea Research Institute For Human Settlements, No. 450, pp. 1-6 (in Korean).
- Jennings, M.E., and Benson, M.A. (1969). "Frequency curves for annual flood series with some zero events or incomplete data." *Water Resources Research*, AGU, Vol. 5, No. 1, pp. 276-280.
- Kedem, B., Chiu, L.S., and Karni, Z. (1990). "An analysis of the threshold method for measuring area-average rainfall." *Journal of Applied Meteorology*, AME, Vol. 29, No. 1, pp. 3-20.
- Kim, K.J., and Yoo, C.S. (2007). "On the variations of spatial correlation structure of rainfall." *Journal of Korea Water Resource Association*, KWRA, Vol. 40, No. 12, pp. 943-956 (in Korean).
- Kim, Y.S., Park, M.J., Kim, S.J., Moon, K.H., and Kim, H.S. (2010). "Estimation of frequency based snowfall depth and maximum snowfall depth in 2010, Korea." *Journal of Korea Water Resource Association*, KWRA, pp. 1476-1480 (in Korean).
- Korean Society of Hazard Mitigation (2012). *Fundamentals of disaster management*. Goomibook (in Korean).
- Kwon, S.H., Park, H.S., and Chung, G.H. (2016). "Analysis of snow vulnerability and adaptation policy for heavy snow." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 16, No. 2, pp. 363-368 (in Korean).
- Lee, J.J., Jung, Y.H., and Lee, S.W. (2007). "A study on the evaluation of probable snowfall depth in Korea." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 7, No. 2, pp. 53-63 (in Korean).
- Ministry of Public Safety and Security, Republic of Korea (2015). *Development of the advanced analytical techniques for the effect of natural disaster considering regional characteristics* (in Korean).
- Ministry of the Interior and Safety (MOIS) (2018). *2017 statistical yearbook of natural disaster* (in Korean).
- Park, H.S., and Chung, G.H. (2019). "Frequency analysis for annual maximum of daily snow accumulations using conditional joint probability distribution." *Journal of Korea Water Resources Association*, KWRA, Vol. 52, No. 9, pp. 627-635 (in Korean).
- Shin, J.Y., Ahn, H.J., Jeong, C.S., and Ahn, J.H. (2018). "Frequency analysis of annual maximum wind speed in Korea using mixture distribution." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 18, No. 7, pp. 61-69 (in Korean).

- Shin, J.Y., and Lee, T.S. (2014). "Parameter estimation of the mixture normal distribution for hydro-meteorological variables using meta-heuristic maximum likelihood." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 14, No. 4, pp. 93-100 (in Korean).
- Shin, J.Y., Kim, T.W., Kim, S.C., and Yoon, S.B. (2013). "Statistical frequency analysis of earthquake data at East Sea using mixed distribution functions." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 13, No. 5, pp. 347-354 (in Korean).
- Strupczewski, W.G., Weglarczyk, S., and Singh, V.P. (2003). "Impulse response of the kinematic diffusion model as a probability distribution of hydrologic samples with zero values." *Journal of Hydrology*, Elsevier, Vol. 27, No. 3-4, pp. 328-351.
- Wang, S.X., and Singh, V.P. (1995). "Frequency estimation for hydrological samples with zero values." *Journal of Water Resources Planning and Management*, ASCE, Vol. 121, No. 1, pp. 98-108.
- Wei, T., and Song, S. (2019). "Probability weighted moments-based parameter estimation for kinematic diffusion and muskingum-based distributions." *Journal of Hydrologic Engineering*, ASCE, Vol. 24, No. 12, pp. 04019054.
- Woo, M.K., and Wu, K. (1989). "Fitting annual floods with zero-flows." *Canadian Water Resources Journal*, Vol. 14, No. 2, pp. 10-16.
- Yoo, C.S., and Kim, K.J. (2004). "Estimation of areal reduction factor using a mixed distribution." *Journal of Korea Water Resource Association*, KWRA, Vol. 37, No. 9, pp. 759-769 (in Korean).
- Yoo, C.S., Jung, K.S., and Kim, T.W. (2005). "Rainfall frequency analysis using a mixed gamma distribution: evaluation of the global warming effect on daily rainfall." *Hydrological Processes*, Wiley, Vol. 19, No. 19, pp. 3851-3861.
- Yu, I.S., Kim, D.H., and Jeong, S.M. (2016). "Assessment of design snow depth for green house using frequency analysis." *The Magazine of The Korean Society of Hazard Mitigation*, KOSHAM, Vol. 16, No. 1, pp. 247-254 (in Korean).