



## Frequency analysis for annual maximum of daily snow accumulations using conditional joint probability distribution

Park, Heeseong<sup>a</sup> · Chung, Gunhui<sup>b\*</sup>

<sup>a</sup>Research Fellow, Department of Land, Water and Environment Research, Korea Institute of Civil engineering and building Technology (KICT), Gyeonggi-do, Korea

<sup>b</sup>Associate Professor, Department of Civil Engineering, Hoseo University, Chungcheongnam-do, Korea

Paper number: 19-039

Received: 5 July 2019; Revised: 29 August 2019; Accepted: 29 August 2019

### Abstract

In Korea, snow damage has been happened in the region with no snowfalls in history. Also, casual damage was caused by heavy snow. Therefore, policy about the Natural Disaster Reduction Comprehensive Plan has been changed to include the mitigation measures of snow damage. However, since heavy snow damage was not frequent, studies on snowfall have not been conducted in different points. The characteristics of snow data commonly are not same to the rainfall data. For example, some parts of the southern coastal areas are snowless during the year, so there is often no values or zero values among the annual maximum daily snow accumulation. The characteristics of this type of data is similar to the censored data. Indeed, Busan observation sites have more than 36% of no data or zero data. Despite of the different characteristics, the frequency analysis for snow data has been implemented according to the procedures for rainfall data. The frequency analysis could be implemented in both way to include the zero data or exclude the zero data. The fitness of both results would not be high enough to represent the real data shape. Therefore, in this study, a methodology for selecting a probability density function was suggested considering the characteristics of snow data in Korea. A method to select probability density function using conditional joint probability distribution was proposed. As a result, fitness from the proposed method was higher than the conventional methods. This shows that the conventional methods (includes 0 or excludes 0) overestimated snow depth. The results of this study can affect the design standards of buildings and also contribute to the establishment of measures to reduce snow damage.

**Keywords:** Censored data, Conditional probability, Design standard of snow loads, Frequency analysis, Snow accumulation

## 적설 자료의 빈도해석을 위한 확률밀도함수 개선 연구

박희성<sup>a</sup> · 정건희<sup>b\*</sup>

<sup>a</sup>한국건설기술연구원 국토보전연구본부 연구위원, <sup>b</sup>호서대학교 건축토목환경공학부 토목공학전공 부교수

### 요 지

우리나라에서도 최근 들어 과거 눈이 내리지 않던 지역에 폭설이 내리거나 대설에 의한 인명피해가 발생하는 등의 설해가 발생하였다. 이에 자연재해저감 종합계획에 대설에 의한 설해 대비를 포함하는 등의 정책적인 변화가 생겼으나, 우리나라는 그동안 대설 피해가 많지 않았었기 때문에 대설이나 적설 자료의 특성에 대한 연구를 다양한 각도에서 수행한 적이 거의 없다. 우리나라의 적설자료는 강우자료와 특성이 다른 경우가 많다. 예를 들어, 우리나라 남해안 일부 지역은 연중 눈이 한 번도 내리지 않는 경우가 다수 있어, 연최대치계열 자료 중에 값이 없는 경우가 빈번히 존재하는 등 중도절단 자료(censored data)와 비슷한 특성을 가진다. 실제로 부산 지점은 적설관측시 시작된 이후 연최대치계열 자료의 값이 없는 경우가 전체 시계열의 36% 이상이었다. 그럼에도 불구하고 적설자료의 빈도해석은 기존 강우자료의 빈도해석 절차에 준해 시행되는 경우가 대부분이었다. 연최대치계열 자료가 존재하지 않는 경우, 기존의 빈도해석 방법을 적용하기 위해 자료가 존재하지 않는 기간에 대해 0으로 가정하여 빈도해석을 수행하거나 해당기간을 제외하고 빈도해석을 수행할 수 있다. 그러나 두 가지 경우 모두 구해진 확률분포의 적합도가 매우 낮은 경우가 존재했다. 그러므로 본 연구에서는 우리나라 적설자료의 특성을 고려하기 위해 조건부결합확률분포를 이용하여 확률밀도함수를 선정하는 방법을 제안하였다. 그 결과 기본 방법에 비해 적합도가 더 높은 확률밀도함수를 구할 수 있었으며, 100년 빈도 이상의 고빈도에서 기존 방법에 비해 대체로 적설심이 작아지는 경향을 보였으며, 최대 15%의 차이를 보였다. 눈의 단위중량에 따라 지역별로 하중은 달라질 수 있으며 그 영향의 크기가 달라질 수 있으나, 본 연구의 결과는 건축물의 설계기준에도 영향을 미칠 수 있고, 재해저감을 위한 대책 수립에도 큰 기여를 할 수 있을 것이다.

**핵심용어:** 빈도해석, 신적설, 적설설계기준, 조건부확률, 중도절단자료

\*Corresponding Author. Tel: +82-41-540-5794

E-mail: gunhuic@gmail.com (G. Chung)

## 1. 서론

2019년 4월 9일 강원도 동해안에 약 20 cm의 폭설이 내려 차량이 눈길에 고립되고 나무가 쓰러지는 등 피해가 속출했다. 강원지방기상청에 따르면, 4월 9일 오전 11시까지 대관령 23.8 cm, 태백 22.5 cm, 평창 용평 21.4 cm, 고성 향로봉 20.6 cm, 강릉 왕산 16.2 cm, 정선 사북 16.6 cm의 눈이 쌓였다고 한다. 2017년 1월 20일에도 강원도 동해안에 약 50 cm의 폭설이 내려 도로의 곳곳이 마비되었다. 최근(2019년) 독일에서는 5월에 갑자기 눈이 내리기도 하였다. 이처럼 이상기후의 징후가 세계 여러 곳에서 나타나는 등 대설에 대한 관심이 증가하고 있다. 이에 행정안전부에서는 2019년 3월 24일 기존 “풍수해저감 종합계획”을 “자연재해저감 종합계획”으로 변경하며, 가뭄과 대설에 대한 대비도 할 수 있도록 하였다. 대설에 대한 대비를 위해서는 대설 위험지역을 미리 선정하여 피해에 대한 대비를 해야 한다. 이를 위해서는 각 지역별로 확률적설심 산정이 필요하다. 과거 우리나라의 적설자료에 대한 확률적설심을 구하기 위해 빈도해석을 수행한 경우를 살펴보면, Yu *et al.* (2014)은 우리나라의 20년 이상 적설심 자료를 보유한 66개 기상관측소의 자료를 빈도 해석하여 연최대치신적설심에 대한 적절한 확률분포형을 제안하였다. 그 결과, 확률가중모멘트법에 의한 Weibull 분포가 가장 적합한 확률분포형이라고 제안하였다. Park *et al.* (2014)은 신경망 모형을 이용하여 RCP 기후변화 시나리오를 적용한 미래 신적설일수와 최심신적설 빈도 분석을 수행하였다. 이를 통해 미래 확률신적설일수와 확률최심신적설심이 대체로 감소하는 경향을 확인하였다.

산정된 확률적설심은 앞서 언급한 바와 같이 재해 저감을 위한 기본 자료로도 활용되지만, 건축물 및 비닐하우스와 같은 시설물의 적설하중 설계기준으로 이용되기도 한다. 예를 들어, 건축구조기준(KBC2016)은 최심적설깊이 자료가 10년 이상 존재하는 38개 지점에 대해 극치확률통계해석을 통해 100년 빈도 적설심을 구하고 이를 지상기본적설하중으로 제시하고 있다. 또한 2010년 농림수산식품부에서 고시된(제2010-128호) 「원예특작시설 내재해형 규격 설계도-시방서」에 따르면, 내재해 설계기준으로 30년 빈도 적설심을 지역별로 제시하고 있다.

반면에 우리나라에서 적설자료에 대한 빈도해석 수행을 위한 자료 수집 방법이나 빈도해석 절차 등에 대한 연구는 그리 많지 않으며, 적설자료의 특성이 강수량 자료와 다소 차이가 있음에도 불구하고 동일한 방법으로 확률적설심을 산정하고 있다. 예를 들면 특정 지점의 경우, 강수량 자료와는 달리 확률적설심 산정에 필요한 연최대치계열 자료의 일부가 누락

되는 현상이 있는데, 이것은 해당 지역에서는 연중 적설이 전혀 없는 상황이 가끔 발생하기 때문이다. 이러한 지점은 특히 남해안 지역에 많이 분포하고 있다. 이러한 경우 확률적설심 산정에 연초과치계열 자료를 이용할 수도 있으나, 이는 연최대치계열 자료에 대한 분석과 다른 결과를 보여 이에 대한 상관관계 등을 따로 분석하여 변환하는 등의 단계를 거쳐야 하므로 통계적 신뢰성이 낮아진다는 단점이 있다. 만약 적설이 발생하지 않은 해의 자료를 0 cm로 하여 연최대치계열의 최심신적설심 자료를 구성하여 빈도분석을 시도할 수도 있으나, 다수의 자료가 0 cm일 경우에는 0 cm인 자료에 의해 분포형의 선택과 매개변수의 추정이 실제 자료와는 다르게 되고, 그 정도가 심할 경우 대부분의 확률분포함수에 대해 적합도 검정이 실패하게 된다.

적설심 자료에 대한 연최대치계열의 이러한 특성은 중도절단자료(censored data)로 알려진 자료의 특성과 매우 유사하다. 중도절단자료란 일정 수준 이하 또는 이상의 값이 관측한계로 인해 특정 임계값으로 기록된 형태의 자료를 말하며, 적설이 발생하지 않은 해를 다수 포함하고 있는 연최대치계열의 적설심 자료 또한 0 미만의 값에 대해 중도절단된 형태를 나타내므로 중도절단에 의한 왜곡을 보정하기 위한 방법을 적설심 자료에도 적용할 수 있을 것으로 판단된다. 이에 본 연구에서는 적설심에 대한 연최대치계열 자료를 중도절단자료로 가정하고, 중도절단자료 보정을 위해 적용할 수 있는 조건부결합확률분포를 적용하여 기존의 방법으로 확률밀도함수를 구한 경우와 그 결과를 비교하였다.

## 2. 중도절단자료(Censored Data)의 확률분포함수 추정

### 2.1 신적설 자료의 특징

빈도해석은 수문관측자료의 수집, 무작위성 검토, 변동 및 경향성 분석, 확률분포형 가정에 따른 매개변수 추정, 적합도 검정 등으로 수행된다. 이 중 양질의 수문자료 관측은 신뢰도 높은 결과를 위해 가장 중요한 절차이다. 자료를 측정하다보면, 특정 기준치 이하의 자료는 관측이 되지 않거나 검출이 불가능하여 특정 임계값(threshold)으로 표현되는 자료들이 존재한다. 이런 자료들을 포함한 시계열 자료를 중도절단자료(censored data set)라고 한다(Cohen, 1991). 이것은 마치 전체 자료들 중에서 어떤 기준 이상이나 이하의 자료를 모두 삭제한 자료들처럼 보이기 때문이다(Kroll and Stedinger, 1996; Maidment, 1993). 예를 들어, 수질 측정시 “최소 검출농도 이하 (less than the detection limit)”로 표현된 자료들을 포함한 자료

계열을 의미한다. 혹은 하천의 최소유량 측정 시, 측정 가능 값 이하인 경우 하천의 유량을 0으로 표현하는 것도 같은 경우이다. 이 경우 하천의 바닥이 실제로 완전히 건천화되어 유량이 0인 경우도 있지만, 측정 가능한 범위와 0 사이 어딘가에 유량 값이 존재하는 경우도 있을 수 있다. 이렇게 측정값이 0인 자료들을 다수 포함하고 있는 시계열 자료의 모멘트 등과 같은 통계적 특성치를 계산하기 위해서는 이 0인 값들을 신중히 다루어야 한다(Maidment, 1993).

중도절단자료를 다루기 위해 probability plot과 probability-plot regression, 가중모멘트 추정, 최우도 추정, 조건부 확률 모형 등을 이용하여 자료의 통계적 특성치를 계산할 수 있다 (Maidment, 1993). David (1981)에 따르면, 중도절단자료는 그 특성에 따라 type I과 type II로 구분할 수 있다. Type I censoring은 특정 값 이하의 값이 모두 censored 되는 경우를 말하며, 이 경우 censored data는 난수발생으로 재현할 수 있다. 이와 달리 type II censoring은 특정수의 자료가 항상 censored 되므로 censored threshold가 무작위로 바뀌는 경우이다. 앞서의 수질 측정과 같은 경우는 사용된 검출 기술에 의한 검출 가능 농도가 일정하므로 type I censoring을 야기한다고 볼 수 있다.

Type I censoring 자료 계열의 통계 특성치 계산을 위해 0으로 관측된 자료를 단순히 0과 검출기준 사이의 무작위 자료로 대체하는 방법이 자주 사용되었다(Cohen and Ryan, 1989; Newman et al., 1989). Gillion and Helsel (1989) and Helsel and Gillion (1986)는 단순히 난수로 자료를 대체하는 것 보다 수질 자료의 분포형을 고려하여 censored data를 대체하는 방법을 적용하여 통계치가 더욱 우수하게 계산되는 것을 보인 바 있다.

**2.2 조건부 확률을 이용한 확률밀도함수 보정**

연최대최심심적설 자료를 살펴보면 연도별로 자료가 있거나 없는 경우와, 만약 있다면 어떤 분포로 있는지를 판단하는 두 단계로 구분하여 생각할 수 있다. 이것은 각 단계별로 다른 확률과정이 적용된다는 것을 의미한다. 이에 본 논문에서는 조건부 확률을 이용하여 다음과 같이 두 가지 확률과정을 조합함으로써 기존 확률밀도함수를 보정하는 식을 유도하였다.

먼저 확률함수  $P(x)$ 에 대해 사건(event)  $B$ 가 발생했을 때, 사건  $A$ 가 발생할 조건부 확률은 다음과 같이 쓸 수 있다.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{when } P(B) > 0. \quad (1)$$

이러한 조건부 확률을 연속확률함수에 적용하여, 특정 조건  $A = \{a \leq X \leq b\}$ 에서 사건  $X$ 가 발생할 확률을 누가한 함수를  $F_{X|A}(x)$ 라고 하면  $F_{X|A}(x)$ 는 다음과 같이 나타낼 수 있다.

$$F_{X|A}(x) = P(X \leq x|A) = P(X \leq x|a \leq X \leq b) = \frac{P(X \leq x, a \leq X \leq b)}{P(A)} \quad (2)$$

이를 조건이 없는 상황에서의 누가확률함수  $F_X(x)$ 를 이용하여 다시 쓰면 다음과 같다.

$$F_{X|A}(x) = \frac{P(X \leq x, a \leq X \leq b)}{P(A)} = \frac{P(a \leq X \leq x)}{P(A)} = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} \quad (3)$$

위의 Eq. (3)은 0에서 1까지 단조증가하는 함수로 제한된 범위에서만 성립하므로  $x$ 의 구간에 따라 다시 쓰면 다음과 같다.

$$F_{X|A}(x) = \begin{cases} 1 & , b < x \\ \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} & , a \leq x < b \\ 0 & , \text{otherwise} \end{cases} \quad (4)$$

$a = 0$ 일 때,  $F_X(a) = F_0$ 라고 하고,  $b \rightarrow \infty$ 이면,  $F_X(b) = 1$ 이 되므로, 상기 조건부 확률은 다음과 같이 표현할 수 있다.

$$F_{X|A}(x) = \begin{cases} \frac{F_X(x) - F_0}{1 - F_0} & , 0 \leq x < \infty \\ 0 & , \text{otherwise} \end{cases} \quad (5)$$

위의 Eq. (5)를 이용하여 전체 누가확률분포함수를 조건부 확률을 포함한 누가확률분포함수로 나타내면 다음과 같다.

$$F_X(x) = \begin{cases} F_0 + F_{X|A}(x) \cdot (1 - F_0) & , 0 \leq x < \infty \\ 0 & , \text{otherwise} \end{cases} \quad (6)$$

또한 누가확률밀도함수  $F_X(x)$ 의 역함수인 분위함수(Quantile function)는 합성함수의 역함수를 이용해 계산할 수 있으며, 역함수를 구하면 다음과 같다.

$$F_X^{-1}(f) = \begin{cases} F_{X|A}^{-1}\left(\frac{f - F_0}{1 - F_0}\right) & , F_0 \leq f < 1 \\ 0 & , \text{otherwise} \end{cases} \quad (7)$$

여기서  $f$ 는 누가확률을 의미한다.

이상과 같이 전체누가확률밀도 함수를 조건부 확률에 대한 누가확률밀도함수로 나타내고 이에 대한 역함수를 구할 수 있음을 보였다.

### 3. 적설빈도해석

#### 3.1 자료수집

우리나라에서 적설심을 측정하는 관측소는 2019년 현재 102개로 파악되고 있으나, 관측소들이 최근에 신설되어 10년 미만의 자료를 보유하고 있는 경우가 상당수 있다. Fig. 1은 102개의 관측소의 자료 보유 기간을 표시한 히스토그램이다. 그림 (a)는 관측소별 관측기간에 대한 히스토그램인데 적설심의 경우 연중 한 번도 적설이 발생하지 않은 경우도 있어 실제 적설이 발생한 해만 고려하여 자료기간을 나타내면 그림 (b)와 같다. 그림 (b)를 살펴보면 서울, 인천 등과 같이 80년 이상의 적설심 자료를 보유하고 있는 관측소도 있으나 대부분

의 관측소가 30년이 약간 넘는 자료를 보유하고 있으며, 10년 미만의 자료를 보유하고 있는 관측소들도 다수 있다.

일반적으로 통계 분석 시 자료에 대한 통계적인 유의성을 확보하기 위한 최소 자료의 수는 대략 30개 전후로 알려져 있다. 이에 본 연구에서는 적설이 없는 연도를 포함하여 30년 이상의 자료가 존재하고, 적설이 없는 연도를 제외하고도 적설심 자료가 25년 이상 존재하는 관측소를 분석 대상으로 하였다. 적설심을 측정하는 우리나라 전체 102개 기상관측소 중 연 최대 최심신적설심 자료를 25년 이상 보유하고 있는 관측소는 총 63개소이며, 관측소의 위치는 Fig. 2와 같다. 이 중 어떤 관측소들은 특정기간 연 최대 적설심 자료가 없거나 0 cm로 기록된 경우를 다수 포함하고 있다. 예를 들어, 부산 관측소는

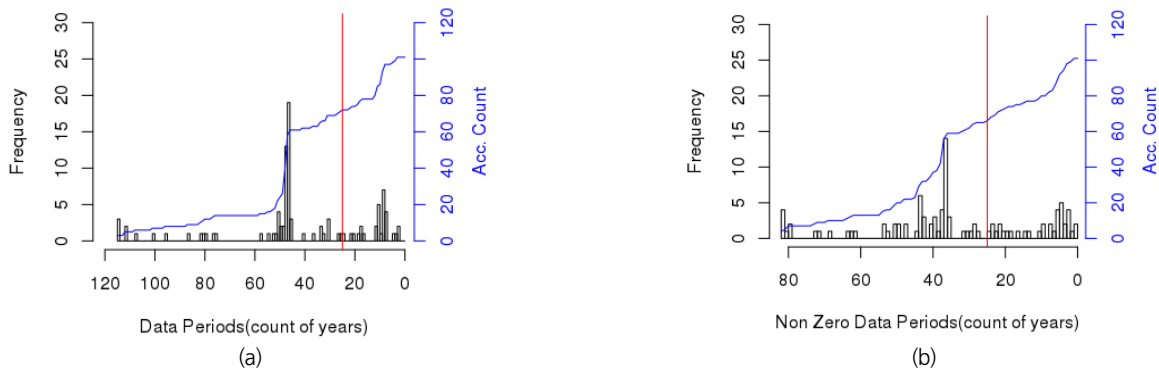


Fig. 1. Frequency of data periods in observatories

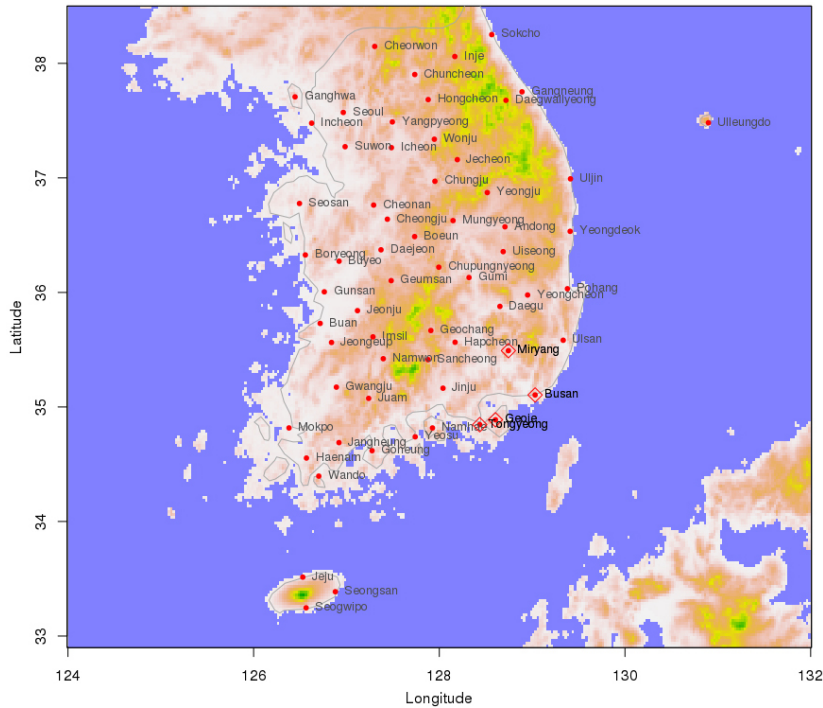


Fig. 2. Locations of observatories

1904년부터 관측을 시작하여 총 115년의 자료가 있어야 하지만 그중 64개년의 적설심 자료가 없거나 0 cm로 기록되어 있으며, 인천도 1904년부터 관측을 시작하여 총 115년 자료 중 43개년의 자료가 존재하지 않거나 0 cm로 기록되어 있다. 자료가 존재하지 않거나 0 cm로 기록된 경우를 보다 자세히 살펴보면, 특정시기 이전 자료가 존재하지 않는 경우가 있다. 예를 들어, 부산은 1939년, 인천은 1949년 이전 적설관측자료가 모두 존재하지 않는다. 그러므로 실제 기상관측을 시작한 시기와 적설관측을 시작한 시기가 일치하지 않음을 알 수 있다. 이에 부수적인 자료 검토 결과, 부산은 1939년, 인천은 1949년부터 적설자료가 유효한 것으로 판단하였다.

이렇게 기상관측과 적설관측 시점의 불일치는 전국의 관측소에 걸쳐 나타난다. 예를 들어, 강릉관측소의 108개년 자료 중 35개년의 자료가 누락되어 있다. 이런 경우, 실제로 연중 적설이 전혀 없었던 것인지, 자료의 단순 누락인지를 구별하는 것이 쉽지 않다. 자료의 단순 누락인 경우는 확률밀도함수를 구할 때 본 연구에서 제시한 조건부 확률을 적용하는 것과는 다른 방법을 적용해야 한다.

본 연구에서는 각 지점별 자료를 검토하여 남해안 지역을 중심으로 적설관측 시점 이후 연중 적설이 없었거나 적설심이 0으로 기록된 경우가 다수 발생한 부산, 통영, 밀양, 거제의 4개 지점을 중점 연구대상지점으로 선택하였으며, 해당지점은 Fig. 2에 가운데 점이 있는 마름모로 강조해 표시하였다. 그리고 Fig. 3에는 각 지점의 연최대최심신적설의 연도별 분포를 나타내었다.

Table 1에서는 각 지점의 기상관측시작일과 유효한 자료 연수, 적설이 없는 해의 수, 그 비율 등을 요약하였다. 기상관측 시작일이 적설관측시작연도와 일치하지 않아 유효한 자료 연수는 기상관측연수에 비해 더 적었다. 자료에 나타난 바와 같이, 부산, 통영, 밀양, 거제 지점은 자료 보유기간 중 적개는 21.6%부터 많게는 36.2%까지 연최대최심신적설이 없는 지점들이다. 그러므로 확률밀도함수를 추정할 때, 이를 0으로

Table 1에서는 각 지점의 기상관측시작일과 유효한 자료 연수, 적설이 없는 해의 수, 그 비율 등을 요약하였다. 기상관측 시작일이 적설관측시작연도와 일치하지 않아 유효한 자료 연수는 기상관측연수에 비해 더 적었다. 자료에 나타난 바와 같이, 부산, 통영, 밀양, 거제 지점은 자료 보유기간 중 적개는 21.6%부터 많게는 36.2%까지 연최대최심신적설이 없는 지점들이다. 그러므로 확률밀도함수를 추정할 때, 이를 0으로

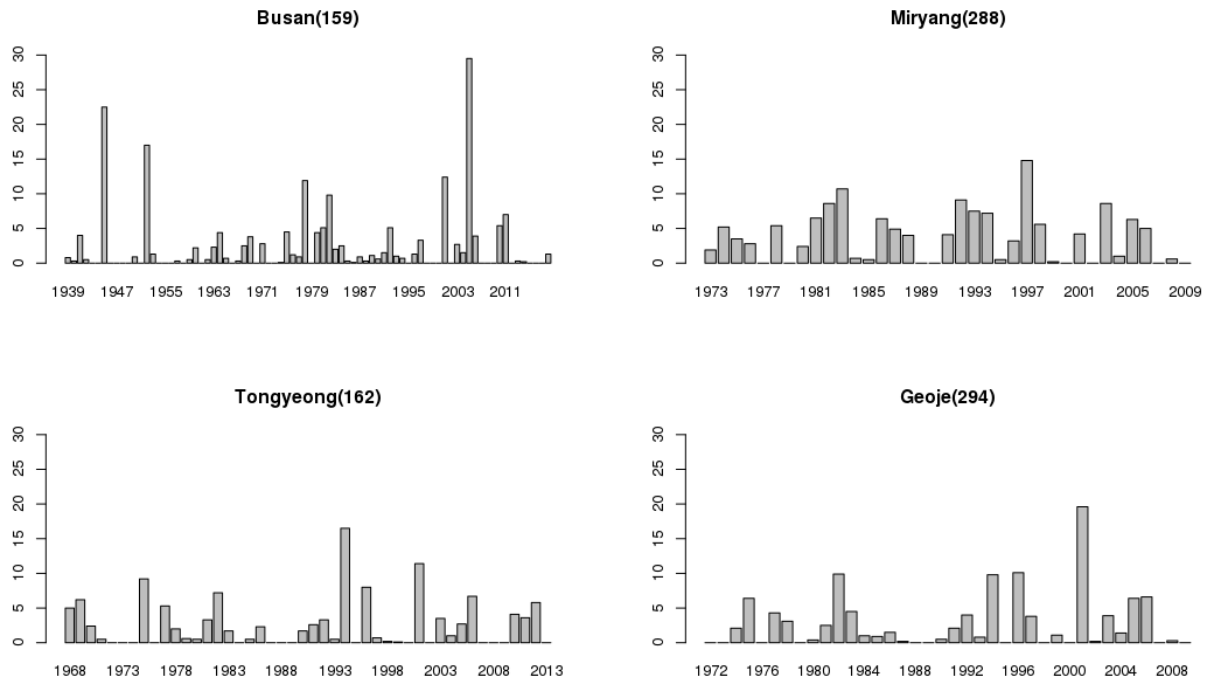


Fig. 3. Time series of annual maximum snowfall on the observatories

Table 1. Selected weather stations and number of zero data

No	Name	Observation Starting Date	Data Period (year)	No. of 0 cm (year)	% of 0 cm (%)
159	Busan	1904-04-09	80	29	36.25
162	Tongyeong	1968-01-01	46	15	32.60
288	Miryang	1973-01-01	37	8	21.62
294	Geoje	1972-01-24	38	11	28.95

고려하면 다수의 0 자료로 인해 적합도 검정 시 귀무가설을 기각해야하는 상황이 발생할 수 있다. 이것은 알려진 이상적인 확률분포형들은 다수의 0이 포함된 자료의 분포와 일치할 수 없는 모양을 갖고 있기 때문이다.

**3.2 확률분포형 결정**

본 연구에서는 연최대최심신적설의 확률분포형으로 Table 2에 정리한 것과 같이 Generalized normal (GNO) 분포형, Generalized extreme-value (GEV) 분포형, Generalized logistic (GLO) 분포형, Pearson Type III (PE3) 분포형, Generalized pareto (GPA) 분포형의 5개 확률분포형을 고려하였으며, 적절한 확률분포형의 선택은 표본의 L-모멘트를 이용하여 구한 매개변수를 적용해 구한 이론분포와 표본분포에 대해 k-s 검정 (Kolmogorov-Smirnov Test)을 실시하여 귀무가설이 채택된 확률분포형을 선택하였다. 이 때 채택된 확률분포형이 다수인 경우에는 유의확률이 가장 큰 경우를 선택하였다.

Fig. 4에서는 선택된 4개의 지점을 대상으로 1) 기존의 방법대로 0을 포함하여 확률분포를 선택하는 방법(적색)과 2) 0을 제외하고 확률분포를 선택하는 방법(녹색), 3) 조건부결합확률분포를 이용하여 0을 고려하는 방법(청색)으로 구분하여 확률밀도함수를 구한 결과를 Q-Q plot의 형태로 나타냈으며, 선택된 확률분포형에 대한 유의확률(p)을 범례에 표시하였다.

자료의 0을 포함하여 분석한 경우와 포함하지 않고 분석한 경우, 각각에 대해 선택된 확률분포형은 Table 3에 정리하였다.

Fig. 4를 자세히 살펴보면, 첫 번째 방법인 0을 포함하여 확률분포형을 선택한 경우(적색선)는 0을 포함한 자료(적색원)를 표현하기 위해 짧은 재현기간에서 기울기가 낮아지고 그 영향으로 긴 재현기간의 기울기가 급격히 높아짐을 알 수 있다. 특히, 통영, 거제에서는 50년 이상 긴 재현기간의 적설심이 상대적으로 크게 산정된다. 반대로 밀양의 경우에는 선택된 확률분포형이 0인 구간을 제대로 표현하기 어려운 분포형이 선택되었으며, 100년 이상 긴 재현기간에서 오히려 값이 작아지는 경향을 보인다.

두 번째 방법은 0을 제외하고 확률분포형(녹색선)을 추정한 경우이다. 0을 제외한 자료(녹색원)에 적합한 확률분포형을 추정하였으므로, 제외된 자료수 만큼 재현기간이 짧아져 0을 포함한 자료에 비해 확률밀도함수가 왼쪽으로 편이되어 있는 것을 확인할 수 있다. 그러므로 일반적으로 긴 재현기간으로 갈수록 적설심이 과대추정 된다. 하지만 첫 번째 방법에 비해 유의확률이 큰 확률분포형을 얻을 수 있다는 점이 장점이다.

마지막으로 조건부결합확률분포를 이용하여 구한 확률분포형(청색선)을 0을 포함한 자료(적색원)와 비교하면, 0이 포함된 자료를 잘 표현하면서도 긴 재현기간에서 적설심을 적정하게 추정하고 있는 것을 알 수 있다. 대부분의 경우 0을 제외하

**Table 2.** Probability distributions for this work

Distribution	pdf and/or cdf	Note
Generalized normal (GNO)	$F(x) = \Phi(y)$	$y = -\frac{1}{\kappa} \log\left(1 - \frac{\kappa(x-\xi)}{\alpha}\right)$
Generalized extreme-value (GEV)	$F(x) = e^{-e^{-y}}$	
Generalized logistic (GLO)	$F(x) = \frac{1}{1+e^{-y}}$	
Generalized pareto (GPA)	$F(x) = 1 - e^{-y}$	
Pearson Type III (PE3)	$f(x) =  x - \xi ^{(\alpha-1)} \frac{e^{-\frac{ x-\xi }{\beta}}}{\beta^\alpha \Gamma(\alpha)}$	

$\Phi(y)$  is the distribution function of the standard normal distribution  
 $\Gamma(\alpha)$  is the gamma function

**Table 3.** Selected type of appropriate probability distribution for frequency analysis

Station	Code	Selected PDF when 0s are included	Selected PDF with conditional probability
Busan	159	GNO	GNO
Tongyeong	162	GPA	GPA
Miryang	288	GPA	PE3
Geoje	294	GNO	PE3

\* GNO: Generalized normal distribution, GEV: Generalized extreme-value distribution, GLO: Generalized logistic distribution, PE3: Pearson type III distribution GPA: Generalized pareto distribution.

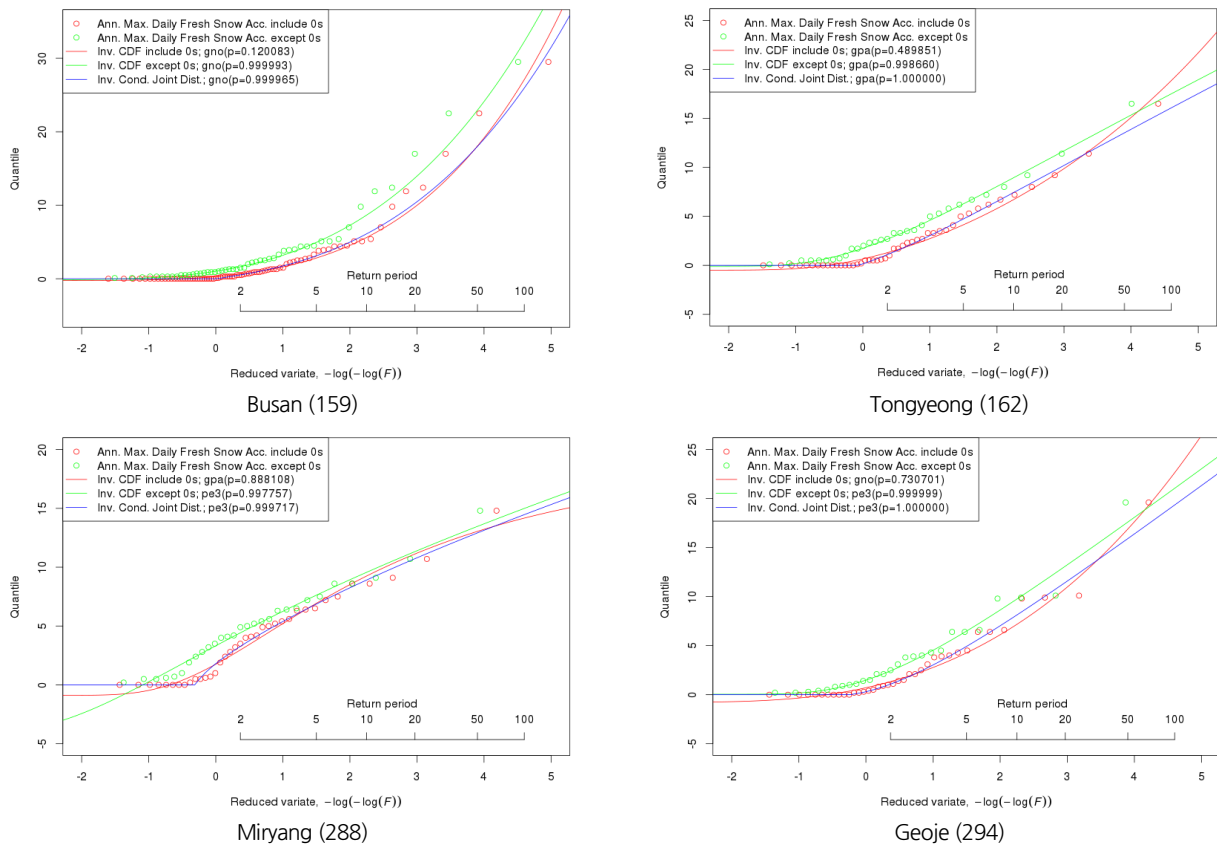


Fig. 4. Q-Q plots functions with/without zero data and conditional probability

고 추정된 확률밀도함수를 오른쪽으로 이동시켜 놓은 것 같은 형태를 보이고 있는데 이것은 조건부결합확률의 효과이다.

조건부결합확률 방법은 적설의 유무에 대한 확률과 적설이 있는 경우 적설심의 확률분포를 합성한 것이라고 볼 수 있는데, 이때 적설이 있는 경우에 해당하는 확률부분을 0을 제외하고 추정된 확률밀도함수가 대체하는 형태를 갖고 있기 때문이다. 이에 0을 제외한 자료들의 확률밀도함수와 조건부결합확률 분포의 형태가 비슷하고 오른쪽으로 0이 발생하는 확률에 의해 이동하는 형태를 보인다. 특히 조건부결합확률분포를 이용한 경우 밀양과 같이 0과 0을 제외한 최소적설심 사이의 간격이 크거나 저빈도 구간에서 적설심의 크기가 급격히 변하는 경우에도 자료의 형태를 잘 표현하는 것으로 나타났다.

Fig. 4의 첫 번째 경우인 0을 포함하여 선정된 확률분포형과 세 번째 경우인 조건부결합확률분포를 이용하여 선정된 확률분포형이 다른 경우도 있다. Table 3에 표시된 바와 같이, 부산, 통영은 같은 분포형이 선택되었지만, 밀양, 거제는 선택된 분포형이 달라졌다. 이는 적합도 검정 결과 유의확률이 큰 분포형을 선택했기 때문이다.

### 3.3 확률적설심

Tables 4 and 5는 각 지점별로 두 가지 방법으로 추정된 확률 밀도함수를 이용해 산정한 30년, 100년 빈도의 확률적설심과 그것을 비교한 결과이다. 30년 빈도의 확률적설심은 대체적으로 두 가지 방법의 차이가  $\pm 3\%$  이내인 것으로 나타났지만, 100년 빈도의 경우는 대체적으로 조건부결합확률밀도를 적용했을 때 대략 10% 이상 적설심이 작게 추정되는 것으로 나타났다. 상세히 살펴보면, 거제는 조건부결합확률분포를 적용한 경우, 3.3 cm의 적설심이 작아졌으며, 밀양은 2.8 cm가 작아졌다. 그러나 통영의 경우, 조건부결합확률분포를 적용한 경우 100년 빈도 확률적설심이 0.3 cm가 증가했다. 이는 조건부결합확률밀도를 적용한다고 확률적설심이 일률적으로 감소하거나 증가하는 것은 아니라는 것을 보여준다.

### 3.4 적설하중 비교

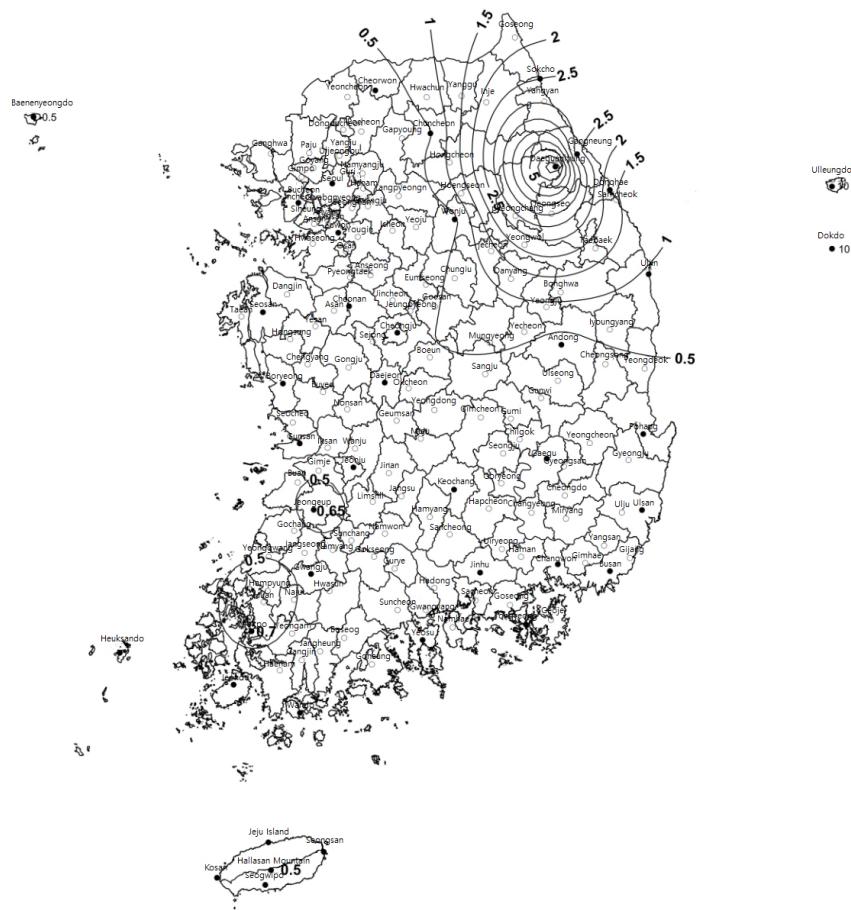
2016년에 발행된 건축구조기준(KBC) (Fig. 5)에서는 100년 빈도 적설심을 이용해 기본지상적설하중을 제안하고 있는데, Fig. 5에서 알 수 있는 것과 같이 남해안 지역은  $0.5 \text{ kN/m}^2$ 을

**Table 4.** 30-years frequency expected snowfall comparison

Station	when 0s are included (cm)	with conditional probability (cm)	difference (cm)	ratio (%)
Busan (159)	13.0	13.3	+0.3	102.7
Tongyeong (162)	11.7	11.6	-0.1	99.1
Miryang (288)	12.0	11.7	-0.3	97.3
Geoje (294)	13.3	13.4	+0.1	100.8

**Table 5.** 100-years frequency expected snowfall comparison

Station	when 0s are included (cm)	with conditional probability (cm)	difference (cm)	ratio (%)
Busan (159)	27.1	25.9	-1.2	95.8
Tongyeong (162)	18.9	16.1	-2.8	85.3
Miryang (288)	14.2	14.5	+0.3	102.1
Geoje (294)	22.6	19.3	-3.3	85.5



**Fig. 5.** Ground Snow Load (KBC, 2016) (Unit:  $\text{kN/m}^2$ )  
 (\*● : stations with historical data, ○ : stations without historical data)

제한하고 있다. 설계용 지붕적설하중은 기본지상적설하중을 기준으로 기본지붕적설하중계수, 노출계수, 온도계수, 중요도 계수 및 지붕의 형상계수와 기타재하분포상태 등을 고려하여 산정하도록 하고 있다. 최소 지상적설하중은  $0.5 \text{ kN/m}^2$ 이다.

Fig. 4에서 과대추정된 약 3 cm를 적설의 단위중량을 고려하여 하중으로 변환해 이와 비교해 볼 수 있다. 적설의 단위중량은 지역, 온도, 습도 등에 따라 달라지는 것으로 알려져 있다. 눈의 밀도와 관련된 연구로 Judson and Doesken (2000)은



Central Rocky Mountain의 11개의 관측소를 대상으로 눈의 밀도를 측정하여 분석하고, 각 관측소별 눈의 특성을 비교하였으며 눈의 밀도와 온도, 지역 등과의 상관관계에 대해서 연구하였다. 그 결과 눈의 밀도는  $10 \sim 257 \text{ kg/m}^3$ 으로 조사되었으며, 대부분은  $60 \sim 100 \text{ kg/m}^3$ 이었다.

이 결과를 적용하면 Table 3에서 과대 추정되었다고 짐작되는 약 3 cm의 100년 빈도 적설심을  $60 \sim 100 \text{ kg/m}^3$ 의 단위중량을 적용하여 하중으로 환산할 경우, 약  $17.64 \sim 29.4 \text{ N/m}^2$ 이므로 건축구조기준의 남해안 지역에 적용된 최소 지상적설하중  $500 \text{ N/m}^2$ 의 3.53~5.88% 정도가 된다. 만약 기온이 그리 낮지 않은 해안가임을 고려한다면 습설일 가능성이 높다. 습설일 경우에는 적설의 단위중량이 두 배로 커지는 점을 고려할 때, 3 cm 적설의 중량은 약  $35.28 \sim 58.80 \text{ N/m}^2$ 으로 건축구조기준의 7.06~11.76%에 이른다. 적설하중이 건축물에 작용하는 다양한 하중들 중 큰 비중을 차지하지 않는 경우가 많을 수 있으나 정확한 적설빈도를 계산하는 것은 방재시설계획을 수립하는 것 뿐만 아니라 건축물의 설계에도 매우 중요한 역할을 할 수 있다.

#### 4. 결론

본 연구에서는 최근 들어 관심이 증가하고 있는 설해와 관련하여 우리나라 적설자료의 특성에 적합한 확률밀도함수를 선정하기 위한 방법론을 제안하고자 하였다. 우리나라 남해안 일부 지역은 연중 눈이 한 번도 내리지 않는 경우가 있으므로, 연 최대치계열 자료의 값이 없거나 0 cm로 기록된 경우가 다수 존재한다. 실제로 어떤 지점은 연 최대치계열 자료의 값이 없는 경우가 전체 시계열의 36%를 넘는 경우도 존재했다. 이런 경우 강우자료의 빈도해석 절차를 준용하여 자료가 없는 해의 값을 0으로 가정해 0을 포함하고 분석을 하는 경우나, 제외하고 분석을 하는 경우 모두에서 확률분포의 적합도가 매우 낮은 결과를 보일 수 있다. 그러므로 본 연구에서는 시계열 자료에 0이 다수 존재하는 지점 4개를 선택하고, 기존 강우자료에 적용하는 빈도해석방법을 적용한 경우와 조건부결합확률을 고려하여 확률밀도함수를 선택하는 경우를 비교하였다. 기존 빈도해석방법을 0인 자료를 포함하는 경우와 제외하는 경우로 나누어 적용하였다. 그 결과 기존보다 적합도가 높은 확률밀도함수를 구할 수 있었으며, 특히 100년 이상의 고 빈도에서 기존 방법과 확률적설심이 달라지는 결과를 보였다. 이는 기존의 방법으로 산정된 확률적설심이 실제자료의 특성을 잘 표현하지 못한 결과일 수 있으며, 향후 추가적인 연구를 통해 달라지는 눈의 단위중량에 따라 설계 하중 등에 미치는 영

향을 계산하고 재해저감 대책 수립에도 큰 기여를 할 수 있을 것이다.

#### 감사의 글

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2017 R1A2B2008828).

#### References

- Cohen, A. C. (1991). *Truncated and censored samples*. Marcel Dekker, New York.
- Cohen, M. A., and Ryan, P. B. (1989). "Observations less than the analytical limit of detection: A new approach." *The Official Journal of the Air & Waste Management Association, JAPCA*, Vol. 39, No. 3, pp. 328-329.
- David, H. A. (1981). *Order statistics*. John Wiley, New York.
- Gillion, R. J., and Helsel, D. R. (1989). "Estimation of distributional parameters for censored trace level water quality data, 1, Estimation techniques." *Water Resources Research*, Vol. 22, No. 2, pp. 135-146.
- Helsel, D. R., and Gillion, R. J. (1986). "Estimation of distributional parameters for censored trace level water quality data, 3, Validation techniques." *Water Resources Research*, Vol. 22, No. 2, pp. 147-155.
- Judson, A., and Doesken, N. (2000). "Density of freshly fallen snow in the central rocky mountains." *Bulletin of the American Meteorological Society*, Vol. 81, No. 7, pp. 1577-1588.
- Korean Building Code (KBC) (2016). *Korean Building Code*. Architectural Institute of Korea, Seoul.
- Kroll, C. N., and Stedinger, J. R. (1996). "Estimation of moments and quantiles using censored data." *Water Resources Research*, Vol. 32, No. 4, pp. 1005-1012.
- Maidment, D. R. (1993). *Handbook of hydrology*. McGraw-Hill, New York, N.Y.
- Newman, M. C., Dixon, P. M., Looney, B. B., and Pinder, J. E. (1989). "Estimating mean and variance for environmental samples with below detection limit observations." *Water Resources Research*, Vol. 25, No. 4, pp. 905-916.
- Park, H. S., Jeong, S. M., and Chung, G. H. (2014). "Frequency analysis of future fresh snow days and maximum fresh snow depth using artificial neural network under climate change scenarios." *Journal of the Korean Society of Hazard Mitigation*, Vol. 14, No. 6, pp. 365-377.
- Yu, I. S., Kim, H. Y., Chung, G. H., and Jeong, S. M. (2014). "Estimation of snowfall frequency and selection of appropriate probability distribution in Korea." *Journal of the Korean Society of Hazard Mitigation*, Vol. 14, No. 4, pp. 101-109.