

## A Stochastic Generation of Snowfall Data and Probabilistic Snowfall Using Multi-site Markov Chain Model

Se Jin Jeung<sup>#</sup>, Woo Suk Han, Byung Sik Kim<sup>+</sup>

Department of Urban & Environmental Disaster Prevention, Kangwon National University, 346, Joongang-ro, Samcheok-si, Gangwon-do, Korea

### Abstract

This paper sought to generate the stochastic simulative snowfall data through a multi-site Markov Chain model, which was constructed from the historical snowfall data collected at major meteorological stations in Gangwon Province in Korea. The need for a multi-site stochastic simulative generation technique was mandated by such a noticeable difference in the weather characteristics between the two regions in Gangwon Province, the Yeongdong and Yongseo regions, divided by the Taebaek Mountain Range. A stochastic simulative generation technique can take into consideration a spatial association of multiple sites. This paper used the Mann-Kendall and autocorrelation function analyses to identify the predisposition and randomness of the data. Then, a multi-site Markov Chain model was used to simulate the data within the period covered by the existing records in Gangwon Province and the simulated data were compared to the actual historical data using a statistical analysis. Based on the comparison, the probabilistic snowfall forecasts were generated for the next 100 years. The multi-site Markov Chain model developed in this paper took into account a sharp distinction between the Yeongdong and Yongseo regions, and was found to be suitable for simulating the snowfall data for architectural purposes.

**Key words:** multi site Markov-chain model, stochastic generation, probable snowfall

### 1. 서론

강원도 지역은 최근 들어 기후변화로 인해 수많은 자연재해가 발생하고 있다. 특히, 강설은 보통 1~2월 중에 발생하며 지역적으로는 12월 또는 3~4월에 폭설이 내리는 지역도 있다. 폭설은 도시에서 도로의 교통소통에 지장을 초래하고, 농촌지역에서는 비닐하우스, 축사 등의 사유시설에 피해를 일으키고 있다. 분석 대상지역

인 강원도 지역은 여름에는 강수량이 많고, 겨울에는 눈이 많이 내리는 기후학적 특성을 지니고 있다. 강원도는 산지가 많이 차지하고 있기 때문에 강설이 발생하게 되면 국지적으로 차이가 많이 나는 것이 특징이다. 이러한 특징으로 인하여 우리나라에 내리는 눈은 특히 공공시설보다는 사유시설에 피해가 집중되고 있으며, 국지적으로는 도로교통이 마비되어 불편을 초래하는 경우가 많이 있다. 하지만 폭설에 의한 피해를 최소화

<sup>#</sup> The 1st author: Se Jin Jeung, Tel. +82-33-570-6838, Fax. +82-33-570-6501, e-mail. [climate@kangwon.ac.kr](mailto:climate@kangwon.ac.kr)

<sup>+</sup> Corresponding author: Byung Sik Kim, Tel. +82-33-570-6819, e-mail. [hydrokbs@kangwon.ac.kr](mailto:hydrokbs@kangwon.ac.kr)

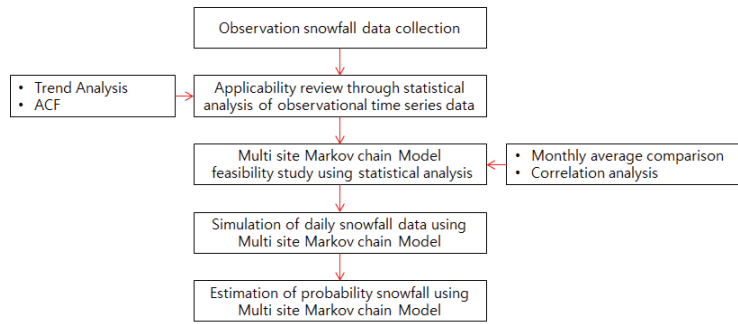


Figure 1. Research flow chart

하기 위한 수문학적 설계기준의 수립이 필요하지만, 관측 강설 자료의 부족으로 인해 강설과 관련된 설계기준 산정 시 어려움이 있다. 따라서 폭설에 의한 피해를 최소화하기 위한 수문학적 계획이 필요하다. 하지만 제한된 강설량자료로 인해 폭설관련 계획을 수립 시 어려움이 있다. 본 논문에서는 부족한 강설량 자료를 추계학적과정을 통해 짧은 자료연한을 갖는 강설량 자료를 장기간의 자료로 모의하여, 설계 강설량을 산정하여 미래에 발생 가능한 폭설재해의 영향 평가를 하고자 한다.

강원도지역은 태백산맥을 중심으로 영서와 영동으로 나누어져 확연히 구별되는 기상특성을 나타내기 때문에 이를 고려 할 수 있는 다 지점 추계학적 모의발생 기법이 필요하다. 본 논문에서는 강원도 주요 기상관측소의 과거 강설자료를 대상으로 다지점 Markov Chain Model을 이용하여 자료를 추계학적으로 모의 발생하였다. 본 논문에서는 먼저 관측치와 동일한 기간의 강설 시계열자료를 모의하여 모형의 매개변수 및 적용성을 검증하였으며 또한, 100년 기간의 강설량 자료로 모의 발생하여 빈도별 설계강설량을 산정하였다.

Markov chain 모형은 주로 일 강수량자료의 모의발생을 위해 이용되는 추계학적 모의발생 기법으로 주로 단일지점 Markov Chain Model이 널리 이용되고 있다. 선행 연구사례로는 Lee & Kim(1995)은 회귀분석과 주성분 분석을 이용하여 결측된 월강수량 자료를 보완하는데 Markov Chain 모형을 이용하였다. Kim(1998)은 다중 상태 Markov Chain을 적용하여 일강수량 자료계열을 모의 발생시켰으며 이를 Tank 모형에 입력 자료로 활용하여 유량 계열에 따른 빈도해석을 실시한

바 있다. Kwon, et. al.(2004)은 기존 댐 재개발의 적정규모 결정을 위하여 비동질성 Markov Chain 모형으로부터 장기간의 일강수량 자료를 모의하여 이용하였다. Jung, et. al.(2005)은 기존 Markov Chain 모형을 이용하여 추정된 일강수량 자료를 NWC-PC 모형의 입력 자료로 활용하여 갑천 유역에 갈수유량에 따른 필요유량을 산정하였다. 또한 Kwon, et. al.(2009)이 비정상성 Markov Chain Model을 이용하여 통계학적 downscaling 기법을 개발하였다.

본 논문에서는 먼저 관측 강설량 자료를 이용하여 ACF분석을 통해 자료의 무작위성을 파악하였다. 그리고 다지점 Markov Chain Model을 이용하여 강원도지역(강릉, 속초, 대관령, 춘천, 철원, 원주)지역의 관측 자료와 같은 기간의 모의자료를 발생시켰으며, 통계적 분석을 통해 모의 강설량자료를 검토하였다. 검토결과를 근거로 하여 100년의 자료를 모의 발생시켜 확률강설량을 산정하고자 하였다.

## II. 이론적 배경

### 1. 다지점(Multi site) Markov-chain 모형

공간적 상관관계를 재현하기 위한 기본적인 가정으로 Wilks(1998)가 제안한 방법을 이용하였다. 이 방법에서 중요한 가정은 임의의 지점  $k$ 와  $l$ 에 대해 강수발생 또는 강수량의 공간적 상관관계는 임의의 날에서 존재하나, 전날( $t - 1$ )과 임의의 날( $t$ ) 사이에 강수발생 또는 강수량의 공간적 상관관계는 없으며, 또한 임의의 날( $t$ )에 대해 강수발생과 강수량의 상관관계가 없다고 가정하는 것이다.

$$Corr[u_t(k), u_t(l)] = Corr[u_t(k), u_t(l)] \neq 0 \quad \text{식 1a}$$

$$Corr[u_t(k), u_{t-1}(l)] = Corr[u_t(k), u_{t-1}(l)] = Corr[u_t(k), u_t(l)] = 0 \quad \text{식 1b}$$

여기서  $u_t(k)$ 는 지점  $k$ 와 시점  $t$ 에서의 강수발생 확률을 나타내고자 균일분포에서 생성된 난수를 의미하며,  $u_t(k)$ 는 마찬가지로 지점  $k$ 와 시점  $t$ 에서의 강수량을 모의하고자 정규분포에서 생성된 난수를 의미한다. 실제 강수발생과는 식 2처럼 나타나진 않지만, 선행연구의 Wilks(1999)의 연구를 보면 식 1의 가정을 이용하여 좋은 모의결과가 보고되었다.

$$\omega(k, l) = corr[\omega_t(k), \omega_t(l)] \quad \text{식 2a}$$

$$\xi(k, l) = corr[X_t(k), X_t(l)] \quad \text{식 2b}$$

각 지점의 관측강수의 이진변수로부터 추정된 결정계수행렬  $[A_0]$ 을 이용하여 다변량 정규분포로부터 난수를 생성하고, 각 지점의 천이확률을 이용하여 강수와무강수

를 이진변수(0 또는 1)로 변환하여 다지점의 공간적 상관관계가 재현된 모의행렬을 생성 할 수 있다. 그러나 이 과정에서 다변량 정규분포로부터 생성된난수는 연속변수인 데 반해, 강수발생의 관측 값은 이진변수라는 문제점이 생긴다. 따라서 다변량 정규분포로부터 관측강수의 각 지점간의 상관관계행렬  $[A_0]$ 가 반영된 연속변수를 생성하여 이진변수로 변환하면, 변환된 이진변수에서는 관측 결정계수 행렬  $[A_0]$ 가 재현되지 않는다. 이러한 문제점을 해결하기 위해 서로 다른 두 지점  $k$ 와  $l$  사이의 연속변수에 대한 결정계수 식 2a와 이진변수에 대한 결정계수 식 2b에 대한 관계를 찾아  $\omega(k, l)$ 을 보정해 주어야만 목표하는  $\xi(k, l)$ 을 재현할 수 있다. <Figure 2>는 다 지점 markov chain 모형을 이용한 일 시계열 자료의 모의 발생 순서를 나타낸 그림이다.

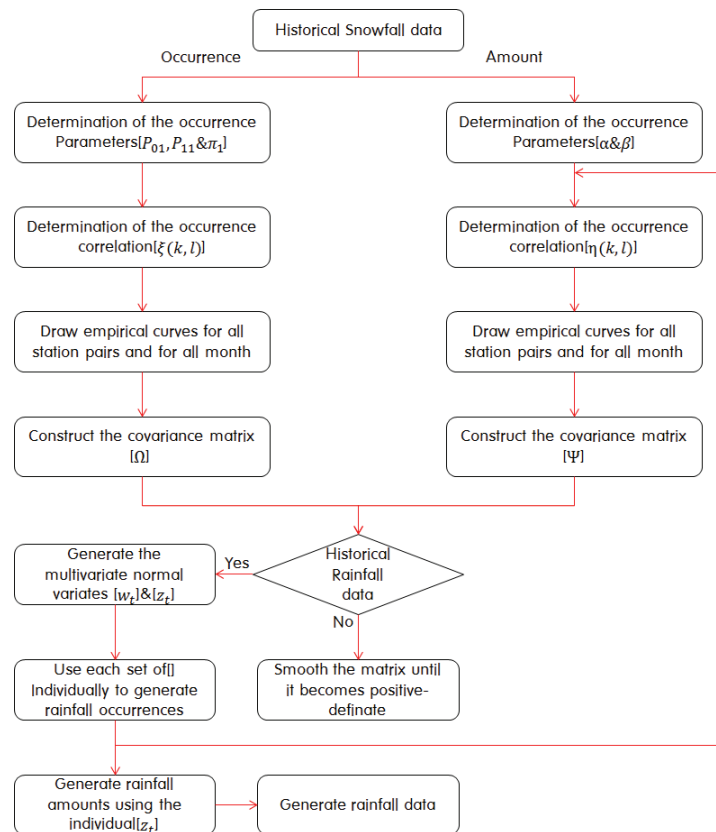


Figure 2. Simulation of snowfall using multi markov chain model

### III. 적용

1. 관측 일강설량 자료의 수집 및 통계학적 분석  
 본 논문에서는 강원도 지역의 주요기상관측소의 일강설량자료를 수집하였으며 추계학적 모형에 적용가능한지를 검토하기 위해 기본 시계열 분석을 실시하였다. 본 논문에서는 우리나라 기후특성을 고려하여 연중 강설량이 가장 많이 발생하는 12월, 1월, 2월, 3월의 자료를 이용하여 분석을 수행하였다. 강원도의 주요관측소의 관측자료의 현황과 위치는 <Figure 3>과 <Table 1>과 같다.

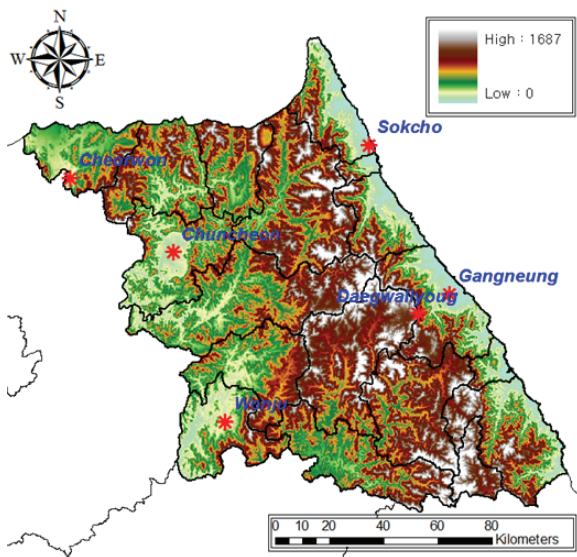


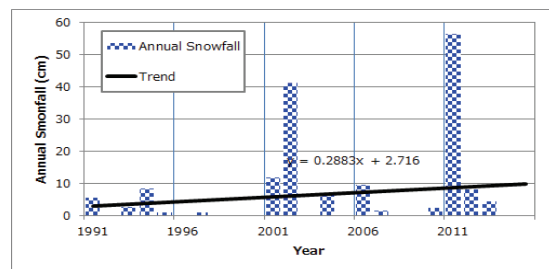
Figure 3. Location map for analysis

Table 1. Target branch observatory status

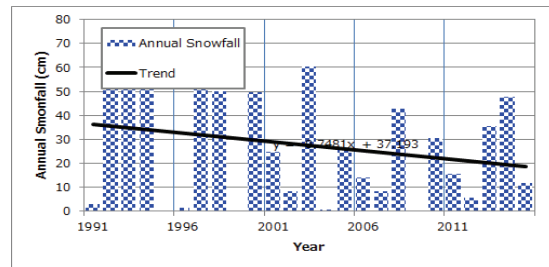
Name	Latitude	Longitude	Data Period
Gangneung	37.75	128.89	1962~2015
Daegwallyoung	37.68	128.72	1971~2015
Sokcho	38.25	128.56	1968~2015
Wonju	37.34	127.95	1972~2015
Cheorwon	38.20	127.25	1988~2015
Chuncheon	37.90	127.74	1966~2015

본 논문에서는 영동과 영서지역의 강설의 발생패턴이 다르다는 것을 입증하기 위해 강설이 발생하는 월(12월, 1월, 2월, 3월)의 경향성 분석을 수행하였으며 ACF분석을 통해 자기상관성분석을 수행하였다. 분석

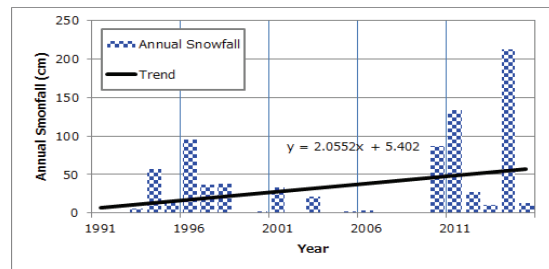
관측소는 영동지역의 강릉, 영서지역의 원주 지역을 선택하였다. 분석결과 강릉의 12월의 경우 증가 경향을 1월의 경우 감소경향, 2월과 3월의 경우 증가 경향을 나타내었다. 원주의 12월의 경우 증가 경향을 1월과 2월은 감소경향을 3월의 경우 증가 경향을 나타내었다 (<Figure 4>, <Figure 5>). 경향성 분석결과 12월 1월 3월은 비슷한 경향을 보였지만 강설이 가장 많이 발생하는 2월의 경우 차이가 발생하는 것을 확인하였다.



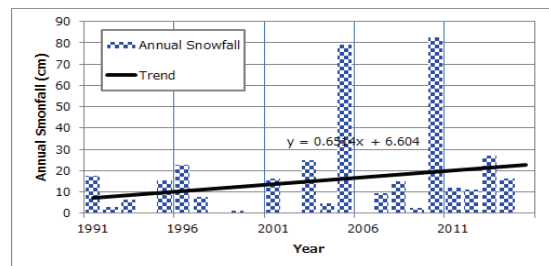
(a) December



(b) January

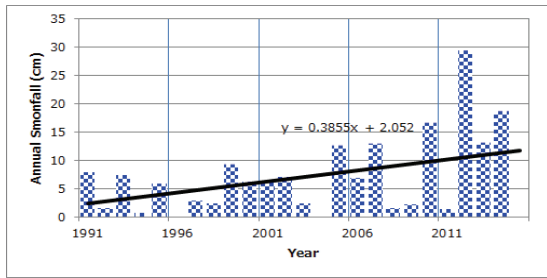


(c) February

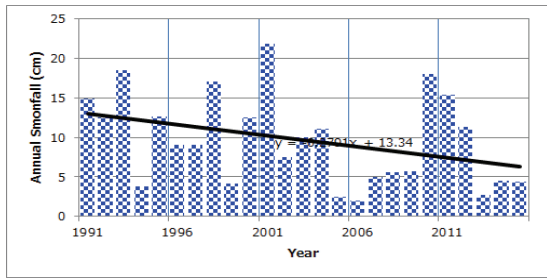


(d) March

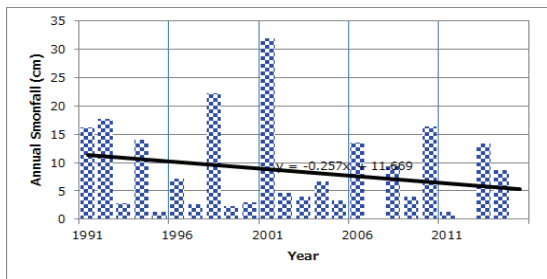
Figure 4. Monthly trend analysis of observed snowfall (Gangneung)



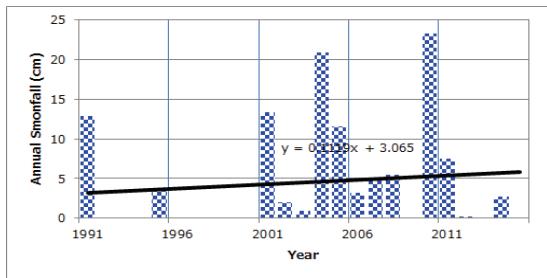
(a) December



(b) January

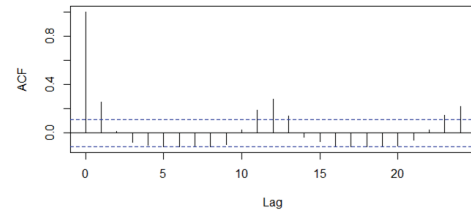


(c) February

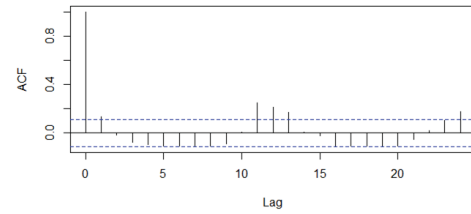


(d) March

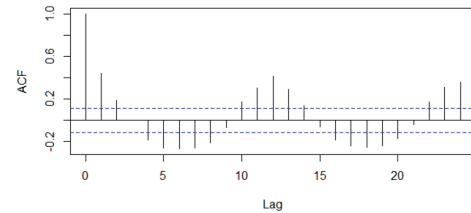
Figure 5. Monthly trend analysis of observed snowfall (Wonju)



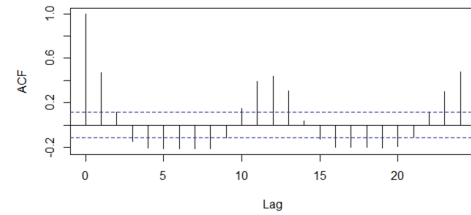
(a) Gangneung



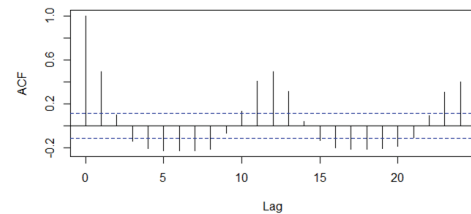
(b) Sokcho



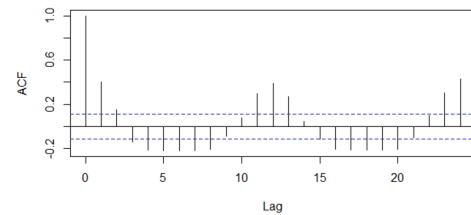
(c) Daegwallyoug



(d) Chuncheon



(e) Cheorwon



(f) Wonju

Figure 6. Results of the analysis of the self-determination coefficient by branch

2. 다지점 Markov Chain Model을 이용한 관측 강설량 시계열자료의 모의

본 논문에서는 다지점 Markov Chain Model 모의자료의 효용성을 확인하기 위하여 강원도의 주요관측소의 관측자료(1990년~2015년, 25년)와 다지점 Markov Chain Model의 모의자료(25년)의 통계치를 검토하였다(Figure 7), (Figure 8)). 본 논문에서는 우리나라 기후특성을 고려하여 년 중 강설량이 가장 많이 발생하는 1월 2월, 3월, 12월의 자료를 이용하여 분석을 수행하였다. (Figure 7)은 관측강설량과 모의강설량의 월 평균을 비교한 것이다. 결과를 살펴보면 강릉의 경우 모의강설량과 관측강설량이 큰 차이가 없었으며, 속초

의 경우 봄철의 관측강설량이 모의강설량보다 많게 분석되었다. 또한 대관령, 춘천, 철원, 원주의 경우 모의강설량이 관측강설량보다 많게 분석되었다. 하지만 6개 관측소모두 큰 차이는 미비하였다. 또한 관측강설량과 모의강설량의 통계분석과 Q-Q plot를 통해 모의강설량 자료의 신뢰성을 검토하였다(Figure 8), (Table 2)). 그 결과 6개 관측소의 상관계수가 0.99로 높은 상관성을 나타내었다. 따라서 앞선 분석결과에 근거하여 볼 때 다지점 Markov Chain Model을 이용한 모의강설량은 장기간의 자료를 생산하는 데 타당성이 있음을 확인하였다.

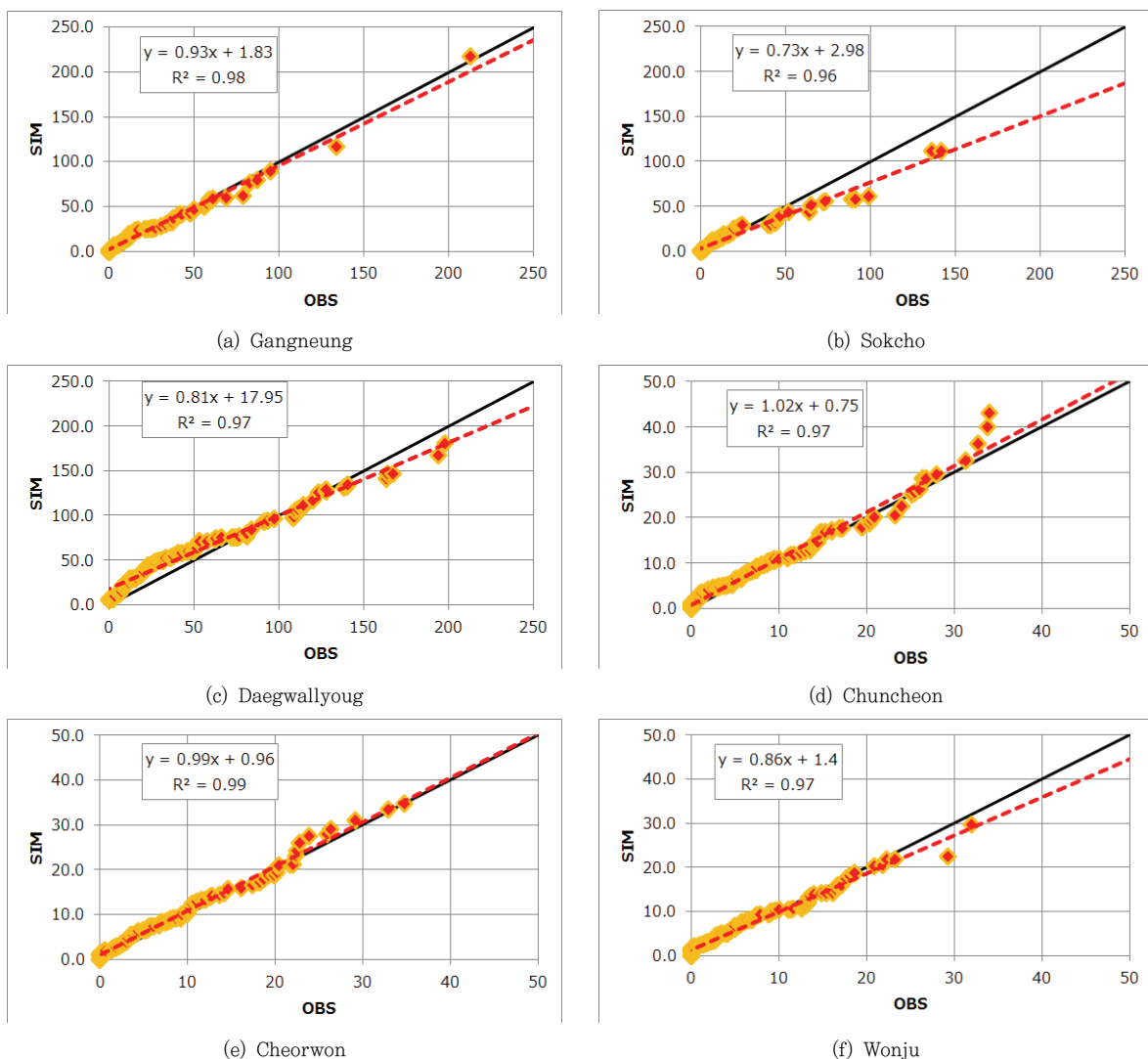


Figure 7. Correlation analysis between observed snowfall and simulated snowfall

Table 2. Statistical analysis of observation data and multi-point markov chain model simulation data

Station	Mean(cm)		Variance		coefficient of determination	Mean error factor (%)
	OBS	SIM	OBS	SIM		
Gangneung	20,3	20,8	32,2	30,2	0,99	2,3
Sokcho	17,6	15,9	28,2	21,1	0,96	-10,6
Daegwallyoug	51,9	60,3	47,8	39,5	0,98	13,9
Chuncheon	9,5	10,4	8,9	9,2	0,98	9,3
Cheorwon	8,8	9,7	8,3	8,3	0,99	9,1
Wonju	7,4	7,8	7,1	6,2	0,98	5,1

3. 다지점 Markov Chain Model을 이용한 일강설량 자료의 확충

앞서 강원도 6개 관측소의 과거 관측 강설량을 다지점 Markov Chain Model에 적용하여 100년치의 모의 일 강설량 자료를 발생시켰으며, 모의 강설량 자료를 이용하여 연강설경향성 분석하였다(Figure 8). 연강설 경향성 분석을 위해 10년의 이동 평균을 계산하였

다. 그 결과 강릉, 속초, 대관령은 각각 10.2cm, 9.4cm, 19.1cm로 평균과의 차가 큰 것으로 분석되었으며, 춘천, 철원, 원주의 경우 각각 3.4cm, 3.6cm, 2.2cm로 평균과의 차가 상대적으로 작은 것을 확인 할 수 있었다. 이는 영동지방과 영서지방의 강설특성이 다르다는 것을 의미한다.

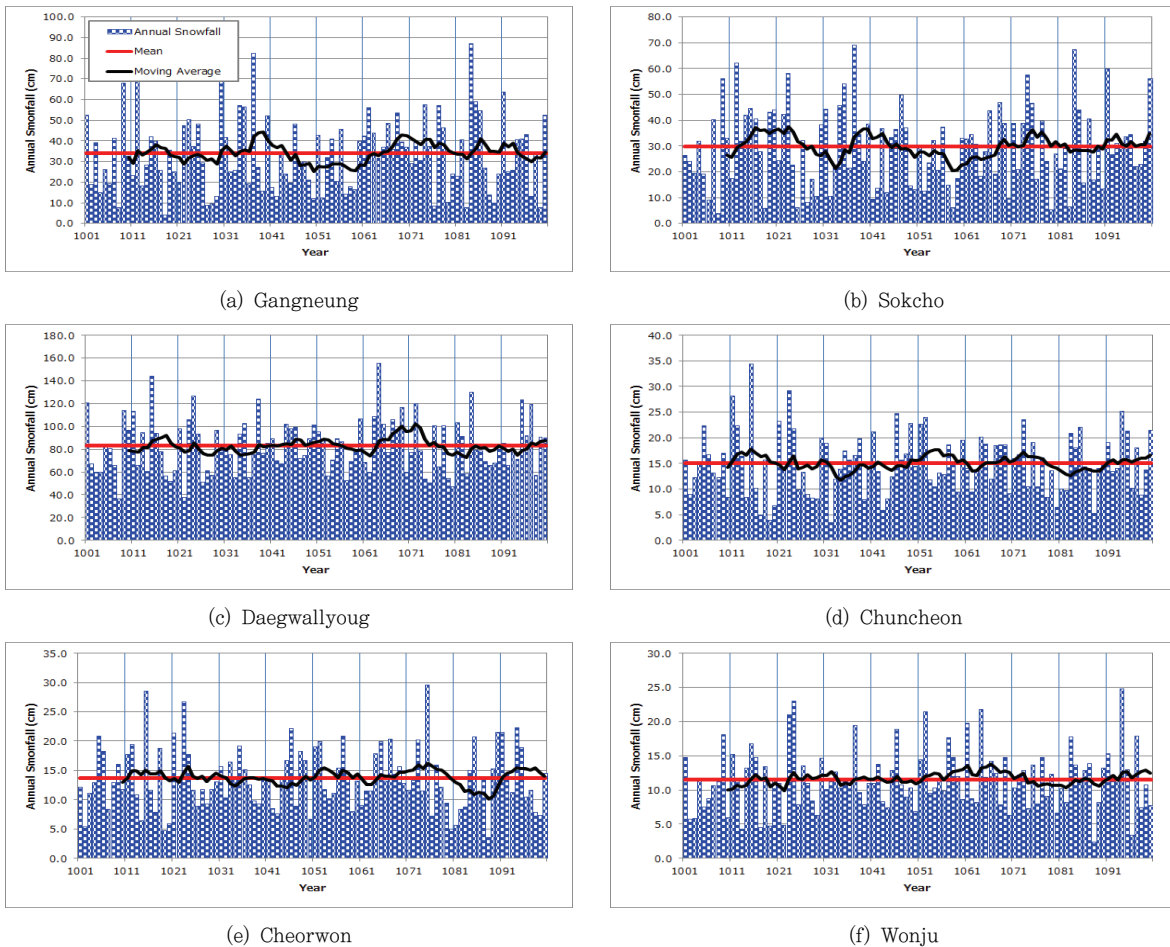


Figure 8. Simulation of snowfall data (100 years)

4. 다지점 Markov Chain Model을 이용한  
확률강설량 산정

다음으로 다지점 Markov Chain Model에서 계산된 각 지점별 100년의 강설량자료의 극한치가 GEV(Generalized Extreme Value) 확률분포형을 따른다고 가정하고, 각 관측소별 연최고치 강설량계열에 대하여 정상성 빈도 분석을 수행하였다. 또한 확률강설량의 지속시간은 1일, 2일, 3일, 4일로 설정하였고, GEV분포형의 매개변

수는 최우도법(Maximum Likelihood)을 사용하여 추정하였다. 각 관측소별 GEV확률 분포형의 매개변수를 추출한 결과 영동지역(강릉, 속초, 대관령)과 영서지역(춘천, 철원, 원주)의 위치매개변수와 규모매개변수가 큰 차이를 보임을 확인하였다. 또한 빈도분석결과 100년 빈도를 살펴보았을 때 영동지역은 평균 32.59cm로 분석되었으며 영서지역의 경우 평균 7.94cm로 분석되었다.

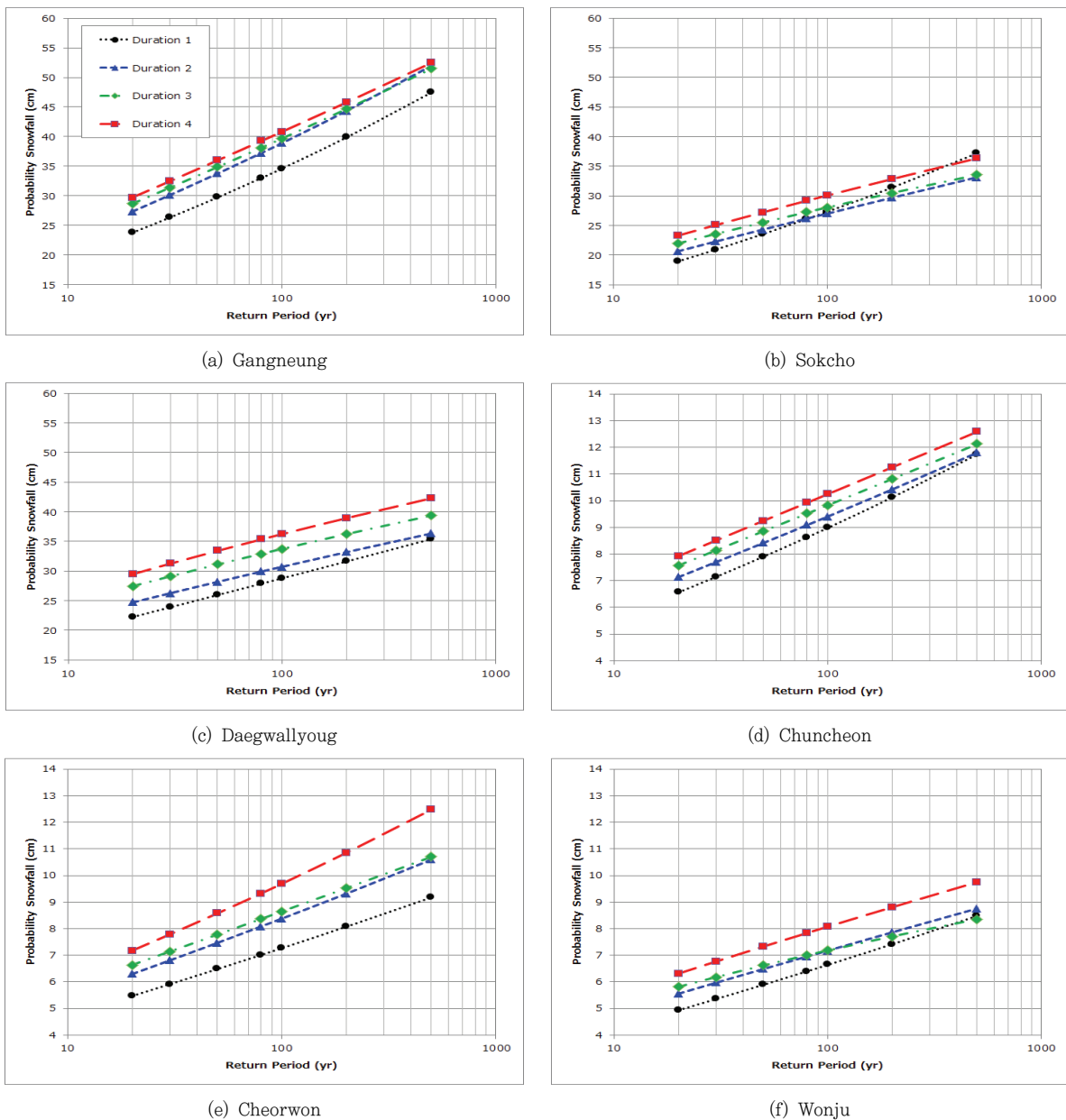


Figure 9. Frequency analysis result

#### IV. 결론 및 향후연구

본 논문에서는 강원도에 위치하고 있는 강릉, 속초, 대관령, 춘천, 철원, 원주의 강설량자료를 이용하여 확률강설량을 산정하고자 했다. 확률강설량을 산정하기 위해서는 최소 30년 이상의 시계열 자료가 필요하다. 하지만 우리나라 강설 시계열 자료는 연속적이지 않고 30년의 충분한 자료가 축적되어 있지 않다. 따라서 본 논문에서는 다지점 Markov Chain Model을 이용하여 강설량자료를 모의발생 시켜 확률강설량을 산정하고자 하였다.

1) 먼저 관측 강설량 자료를 이용하여 경향성분석과 ACF분석을 통해 자료의 무작위성을 파악하였다. 또한 다지점 Markov Chain Model의 모의성능을 확인하기 위해 강원도의 6개 관측소의 관측자료와 다지점 Markov Chain Model을 이용하여 25년의 모의자료를 발생시킨 후 관측자료와 모의자료를 비교 검토 하였다. 검토결과를 요약하면, 강릉의 경우 두 자료 간 큰 차이가 없었으며 속초, 대관령, 춘천, 철원, 원주의 경우 차이가 있었지만, 그 차이는 미미하였다. 또한 Q-Q plot를 통한 신뢰성검토결과 6개 관측소 모두 결정계수가 0.9이상으로 높은 상관성을 나타내었다. 따라서 다지점 Markov Chain Model을 이용한 모의강설량은 장기간의 자료를 생산하는 데 타당성이 있음을 확인하였다.

2) 앞서 언급했듯이 다지점 Markov Chain Model의 모의 강설량은 장기간의 자료를 생산하는 데 타당성이 있음을 확인하였다. 따라서 본 논문에서는 다지점 Markov Chain Model을 이용하여 100년의 강설량 자료를 모의하였다. 각 지점별로 모의된 강설자료가 GEV 확률분포형을 따른다고 가정하고 각 관측소별 연 최고치 강설량 계열에 대하여 빈도분석을 실시하였다. 각 관측소별 GEV확률 분포형의 매개변수를 추출한 결과 영동지역(강릉, 속초, 대관령)과 영서지역(춘천, 철원, 원주)의 위치매개변수와 규모매개변수를 확인 하였을 때 영동지역과 영서지역의 차이가 크게 나타났다. 또한

각 관측소별 지속기간별 재현수준을 확인 해본 결과 강릉, 속초, 철원, 원주의 경우 상대적으로 높은 재현수준을 나타내며, 대관령과 춘천의 경우 낮은 재현수준을 나타내는 것으로 분석되었다.

#### 감사의 글

본 논문은 국토교통과학기술진흥원 지원국토교통기술화사업지원(16TBIP-C111520-01)의 연구비 지원으로 수행되었습니다. 이에 감사드립니다.

#### References

- Jung, Young Hoon, Chung Sung Lee, Hyung Soo Kim, and Myung Pil Shim. 2005. Estimation of Needed Discharge Considering Frequency Based Low Flow in Gabcheon Basin. *Korean Society of Civil Engineers Journal*. 25(2B): 97-105.
- Kang, Kyung Seok. 2000. Simulation of Daily Runoff by Multi Site Day Rainfall Model. Ph.D. Dissertation. Inha University.
- Kwon, Hyun Han and Byung Sik Kim. 2009. Development of Statistical Downscaling Model Using Nonstationary Markov Chain. *Korea Water Resources Association Journal*. 42(3): 213-225.
- Kwon, Hyun Han, Young Il Moon, Byung Kyu Choi, and Yong Nam Yoon. 2004. Optimum Size Analysis for Dam Rehabilitation Using Reliability Analysis. *Korea Water Resources Association Journal*. 38(2): 97-110.
- Lee, Chang Hoon and Seung Kim. 1995. Estimation of Mean Annual and Monthly Precipitations in South Korea by the Regression Analysis. *Korean Society of Civil Engineers Journal*. 1(5): 1255-1266.
- Wilks, D. S. 1998. Multisite Generation of a Dailt Stochastic Precipitation Generation Model. *Journal of Hydrology*. 210: 178-191.
- Wilks, D. S. 1999. Interannual Variability and Extreme-valie Characteristics of Several Stochastic Dailt Precipitation Models. *Agricultural and Forest Meteorology*. 93: 153-169.

*Korean References Translated from the English*

- 강경석. 2000. Multi site 일 강우모형에 의한 일 유출량의 모의 발생. 인하대학교 박사학위논문.
- 권현한, 김병식. 2009. 비정상성 Markov Chain Model을 이용한 통계학적 Downscaling 기법 개발. 한국수자원학회논문집. 42(3): 213-225.
- 권현한, 문영일, 최병규, 윤용남. 2004. 신뢰성 분석을 통한 기존 댐 재개발의 적정규모 결정의 관한 연구. 한국수자원학회논문집. 38(2): 97-110.
- 이창훈, 김승. 1995. 회귀분석에 의한 한국의 연·월평균 강수량의 추정. 대한토목학회논문집. 1(5): 1255-1266.
- 정영훈, 이충성, 김형수, 심명필. 2005. 갑천 유역의빈도별 갈수유량에 따른 필요유량 산정. 대한토목학회논문집. 25(2B): 97-105.

---

Received: Oct. 11, 2017 / Revised: Dec. 6, 2017 / Accepted: Dec. 13, 2017

## 다지점 Markov chain 모형을 이용한 강설량 자료의 추계학적 확충 및 확률 강설량 산정

국문초록 본 논문에서는 강원도 주요 기상관측소의 과거 강설자료를 대상으로 자료를 확충하기 위해 Multi site Markov Chain 모형을 이용하여 자료를 추계학적으로 모의 발생하였다. 강원도지역은 태백산맥을 중심으로 영서와 영동으로 나누어져 확연히 구별되는 기상특성을 나타내기 때문에 이를 고려할 수 있는 Multi site 추계학적 모의발생기법이 필요하다. 그러므로 지역의 특성을 고려할 수 있는 Multi site의 공간적 상관성을 반영할 수 있는 추계학적 모의발생기법이 필요하다. 본 논문에서는 먼저 관측 강설량 자료를 이용하여 ACF분석을 통해 자료의 무작위성을 파악하였다. 그리고 다지점 Markov Chain Model을 이용하여 강원도지역(강릉, 속초, 대관령, 춘천, 철원, 원주)지역의 관측자료와 같은 기간의 모의자료를 발생시켰으며, 통계적 분석을 통해 모의 강설량자료를 검토하였다. 검토결과를 근거로 하여 100년의 자료를 모의 발생시켜 확률강설량을 산정하고자 하였다. 그 결과 영서와 영동지역의 확연히 구분되는 기후특성을 나타낼 수 있는 다지점 Markov Chain 모형을 구성하였으며 이를 통해 설계 강설량 산정에 적합함을 확인 할 수 있었다.

주제어 : 원자력발전, 원자력재난, 원자력사고, 위기관리

Profiles **Se Jin Jeung** : He holds a master's degree in engineering from Kangwon National University. He is a Ph.D candidate of the Department of Urban & Environmental Disaster Prevention Engineering at Kangwon National University. His interesting subject of research is water resources and climate change impact assessment([climate@kangwon.ac.kr](mailto:climate@kangwon.ac.kr)).

**Woo Suk Han** : He is an associate research fellow of the Urban Research Division at Korea Research Institutes for Human Settlement(KRIHS). He received M.S. and Ph.D. Degree in Civil & Environmental Engineering from University of Utah in U.S. His research fields are disaster prevention, climate change, resilience policy etc. He has been the chief research charger of the international joint research between KRIHS and Woodrow Wilson center about disaster prevention resilience for 3 years([wshan@krihs.re.kr](mailto:wshan@krihs.re.kr)).

**Byung Sik Kim** : He received his B.A. M.A., and Ph.D. from Inha University Korea in 2005. He is a professor of the Department of Urban & Environmental Disaster Prevention Engineering at Kangwon National University, in which he has taught since 2011. His interesting subject and area of stochastic hydrology and climate change. He has published about 150 articles in journals and written 3 books ([hydrokbs@kangwon.ac.kr](mailto:hydrokbs@kangwon.ac.kr)).