

바람의 영향에 의한 관측 강우 손실에 대한 베이지안 모형 분석[†]

박영우¹ · 김영민² · 김용구³

¹²³경북대학교 통계학과

접수 2017년 3월 2일, 수정 2017년 5월 9일, 게재확정 2017년 5월 13일

요약

일반적으로 우량계로 측정된 강수량은 지상에 도달한 실제 강수량보다 적게 관측된다. 측정된 강수량이 실제 강수량 보다 적게 측정되는 것은 강수의 형태 (snow, fixed, rain)나 우량계의 종류 그리고 공간적인 특성에 의해 강수량의 정확한 측정이 어렵기 때문이다 (Nitu, 2013). 이는 강수량의 손실을 발생시키는 계통오차 (systematic errors) 때문이며, 일반적으로 고체 강수량의 계통오차는 보통 액체 강수량보다 크다고 알려져 있다. 본 연구에서는 바람에 의한 고체 강수량의 언더캐치 (under-catch)를 알아보고, 겨울에 내리는 모든 강수의 형태 (snow, mixed, rain)에 대하여 연속조정함수를 소개하였다. 이를 위해 고창 표준기상관측소에서 측정된 데이터를 사용하였고, 객관적으로 데이터를 가장 잘 설명하는 모형을 선택하고 평가하기 위해 베이지안 분석을 이용할 것이다. 이번 연구는 강수량 측정에서 Catch Ratio의 계통적 구조에 대한 통계적 분석을 보여주었다.

주요용어: 관측 강우 손실, 베이지안 분석, 연속조정함수, 우량계 자료.

1. 서론

강수량은 지구상의 물에 대해 연구하는 수문모델링과 수증기의 획득, 강수량에 의한 소비 등의 수분 수지 연구에 있어서 중요한 요소이다. 포괄적이고 최적의 수분수지 관찰을 위해서는 정확한 강수량 데이터 이용이 필요하다. 하지만 우량계를 통해서 측정된 강수량은 정확도에 대한 문제가 발생하게 된다. 따라서 세계기상기구 (WMO)는 정확한 기상관측을 위해 일반적인 지침을 제공하며 (WMO, 2008), 강수 측정의 정확성을 높이기 위한 새로운 강수 우량계에 대한 연구 (Strangeways, 2004), 고체형 강수 측정의 정확도를 평가 (Lee, 2014) 등의 정확한 강수량 측정을 위한 연구가 활발하게 이루어지고 있다.

일반적으로 우량계로 측정된 강수량은 지상에 도달한 실제 강수량보다 적게 관측된다. 측정된 강수량이 실제 강수량 보다 적게 측정되는 관측 강우 손실은 강수의 형태 (snow, fixed, rain)나 우량계의 종류 그리고 공간적인 특성에 의해 강수량의 정확한 측정이 어렵기 때문이다 (Nitu, 2013). 이는 강수량의 손실을 발생시키는 계통오차 (systematic errors) 때문이며, 일반적으로 고체 강수량의 계통오차는 보통 액체 강수량보다 크다고 알려져 있다. 계통오차는 수집구 위에서의 바람의 작용, 강수 저장 용기 내부에서의 습윤 손실, 더운 기후에서의 저장 용기로부터 증발, 강수량의 흔적, 기계적 오차, 강수가 수집구의 안팎으로 튀는 현상 그리고 눈이 바람에 날리거나 날려서 쌓이는 오차 등을 포함한다 (WMO, 2008).

[†] 본 논문은 제 1저자인 박영우의 석사학위 논문 중 일부를 정리한 것임.

¹ (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 석사과정.

² (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 조교수.

³ 교신저자: (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 부교수.

E-mail: kim.1252@knu.ac.kr

바람에 의해 강수가 수평적으로 휘날리는 현상 등, 강수량 관측에서 바람은 강수량을 과소 추정하는 가장 큰 환경적 요인이다 (WMO, 2008). 특히, 눈은 비에 비해 바람에 쉽게 잘 휘날리므로 고체형 강수는 액체형 강수보다 풍속에 더 많은 영향을 받는다 (Goodison 등, 1998).

지면에서 적설량과 무게의 실시간 추정은 수문학 예측과 기상 예측에 필요로 되어진다. 이는 하천 저수지 유입 예측에 대해 주요한 정보이고, 수치 예보 모델의 표면변수로서 민감하다고 할 수 있다. 일반적으로, 자동 우량계는 바람이 부는 곳에서 고체 강수량이 과소 추정되는 경향이 있다. 만약 우량계가 제대로 설치되고 올바르게 작동한다면, 이러한 편향은 대부분 우량계의 기체역학과 관련되어 있다. 눈의 조정곡선은 일반적으로 관심 지역에 위치하고 있는 기상관측소에서 검정계이지와 상호비교실험을 통한 통계적 회귀를 이용하여 추정된다.

이 연구의 목표는 바람에 의한 고체 강수량의 언더캐치 (under-catch)를 알아보고, 겨울에 내리는 모든 강수의 형태 (snow, mixed, rain)에 대하여 연속조정함수를 찾아내는 것이다. 이를 위해 고창 표준 기상관측소에서 측정된 데이터를 사용 할 것이며, 객관적으로 데이터를 가장 잘 설명하는 모형을 선택하고 평가하기 위해 베이지안 분석을 이용할 것이다. 이번 연구는 강수량 측정에서 Catch Ratio의 계통적 구조에 대한 통계적 분석을 보여줄 것이다 (Kim과 Kim, 2012; Lee와 Kim, 2016).

2. 자료의 관측장소와 측정방법

2.1. 강수량 관측장소

고창 표준기상관측소는 위도 35°20', 경도 126°35', 고도 52.0m에 위치한다. 모든 기상관측 장비들은 대략 7000m²의 평평한 대지에 설치되어 있다. 기상 관측장소는 겨울철 시베리아 고기압의 영향력이 확대되는 북서풍에 의해 12월에서 1월까지 많은 눈이 관측되는 곳이다. 관측 시설은 강수량과 적설량의 상호비교를 위해 고안되었고, 종관기상관측장비 (automated surface observing system; ASOS)는 관측 장소의 남동쪽에서 위치하고 있다. SPICE (the solid precipitation intercomparison experiment) 프로그램에서, 바람의 영향을 최소화하기 위해 DFIR (double fence intercomparison reference)안에 Tretykov 바람막이를 설치한 자동우량계인 Pluvio400를 검정우량계로 이용한다 (WMO, 2008). 다른 자동우량계들은 DFIR이 설치되지 않은 관측 장소에 위치하고 있으며 144m²에 설치되어 있다. Table 2.1은 고창 관측 지역 자동우량계의 구성, 바람막이와 수집구의 높이를 요약정리 하였다. 바람막이로는 Single alter, Tretykov, PIT, DFIR이 설치되었고, 각각의 자동우량계는 다른 높이로 설치되어 있다. 눈이 많이 발생하고 종종 강풍이 동반되는 장소에서는 고체형 강수에 대한 바람의 영향을 연구하는데 있어 특이한 데이터들을 가져다준다.

Table 2.1 The configurations of automatic precipitation gauge at Gochang standard weather station

Guage	Wind shield	Height of gauge orifice	Number
GEO T200B	Single alter	1.50	1
GEO T200M-3	Single alter	1.50	1
GEO T200M-3	Tretyakov	1.70	1
GEO T200M-3	PIT	0.00	1
PLU 200	Tretyakov	1.70	1
PLU 400	Tretyakov/DFIR	3.00	1
TRws 203	Single alter	0.97	1
TRws 503	Single alter	0.97	1
WPG-A1	Tretyakov	1.00	2
TBRG	No	0.50	2

2.2. 강수량 측정방법과 데이터 설명

고창 표준기상관측소의 Geonor 데이터는 매 10초마다 수집하고, 수집된 수준 값을 더함으로써 1분 강수량 측정 데이터가 얻어진다. Pluvio 우량계는 누적 강수량과 강수량의 강도를 매 1분마다 수집한다. TRwS는 강우율을 측정하고, TBRG는 1분마다 누적강수량을 측정하며 누적강수량의 차이를 이용하여 1분 강수량을 수집한다. Geonor 우량계와 Pluvio 우량계는 미국해양대기관리처 (NOAA)에서 제안된 알고리즘을 통하여 강수량을 측정하고, 우량계의 일 누적 강수량이 3mm보다 크면 강수량은 기록되어진다.

본 논문에서 강수량 자료는 2012년 12월에서 2013년 2월까지 고창 표준기상관측소에서 수집된 자료를 이용하였다. 고체 강수량의 측정을 위한 검정우량계로는 바람의 영향을 최소화 하기 위해 DFIR이 설치되어 있는 우량계를 이용하였으며 (WMO, 2008), 강수량 측정을 위한 자동우량계로는 GEO T200B가 이용되었다. 정확한 강수량이 측정 되었는지 알아보기 위한 척도로서는 Catch Ratio가 사용되었다. Catch Ratio는 검정우량계로 측정된 1시간 누적 강수량과 다른 우량계로 측정된 1시간 누적 강수량의 비로써 정의되는데, P_T 를 검정우량계로 측정된 실제 강수량이라 하고 P_M 을 관측된 강수량이라 하면 Catch Ratio (R)는 다음과 같이 정의된다.

$$R = \frac{P_M}{P_T}.$$

Catch Ratio는 1시간 동안 측정된 검정우량계의 누적 강수량이 0.1mm보다 크면 계산된다. 비를 측정할 때는 수집구를 통해 들어온 비를 모아서 측정을 하는데 원래는 들어와야 하는 강수가 바람에 의해 수집구 밖으로 떨어지게 되므로 Catch Ratio는 원칙적으로는 0~1의 값을 가지게 된다. 만약 얻어진 Catch Ratio가 0~2범위를 벗어난다면, 이번 연구에서 제거하고 분석을 진행하였다. 이 연구에서는 각각의 강수 형태에 따라 바람에 의한 강수 손실에 대한 조정함수를 만들기 위해서 ASOS로부터 제공되는 기온과 지상 10m에서 측정되는 풍속 자료를 이용한다. 일반적으로 강수의 형태에 따라 바람에 의한 강수손실이 다르기 때문에 강수량을 측정할 때 강수의 형태를 구분하여야 한다. 하지만 강수량을 측정할 때 사람이 일일이 확인하기 어려우며 구분 또한 모호하기 때문에 기상청에서는 기온이 영하 2도보다 낮으면 강수의 형태를 눈, 영상 3도보다 높으면 비, 그 사이의 온도에서는 진눈깨비로 다음과 같이 구분할 수 있으며, 구분된 강수의 형태에 따라 각각 모형에 적합시켰다.

$$\text{Precipitation type} = \begin{cases} \text{snow,} & T \leq -2.0^\circ\text{C}, \\ \text{mixed,} & -2.0^\circ\text{C} < T \leq 3.0^\circ\text{C}, \\ \text{rain,} & 3.0^\circ\text{C} < T. \end{cases}$$

3. Catch Ratio에 대한 통계적 모형

유사한 데이터의 특성연구 (Rasmussen 등, 2012)와 강수우량계에서 바람 장에 대한 유체 메커니컬 연구 (Theriault, 2012)를 참고하여 다음의 4가지를 가정 할 수 있다.

1. Catch Ratio는 풍속과 온도의 함수이다.
2. Catch Ratio는 풍속이 0일 때 1에서부터 한계이상까지 단조감소하고, 풍속이 무한의 값을 가지게 되면 0의 값을 가진다.
3. Catch Ratio는 기온에 영향을 받는 풍속의 함수로서 기하급수적으로 감소한다.
4. Catch Ratio는 풍속에 의존하여 두드러지게 변화하고, Catch Ratio는 정의역의 어떤 범위에서는 0값을 가질 수 있다.

Catch Ratio의 모수적 함수에 타당할 것 같은 특징을 고려하여, 수정된 시그모이드 함수를 이용하여 모형에 온도를 포함하였다. Catch Ratio는 0에서 1 사이의 값을 가지기 때문에 수정된 시그모이드 함수를 고려하였으며, 온도 T 와 풍속 v 에 의한 다음의 모수모형으로 표현 할 수 있다 (Wolff 등, 2015).

$$R(T, v; \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau) = \left\{ 1 - \tau_1 - (\tau_2 - \tau_1) \frac{\exp[(T - T_\tau)/s_\tau]}{1 + \exp[(T - T_\tau)/s_\tau]} \right\} \exp[-(v/\theta)^\beta] + \tau_1 + (\tau_2 - \tau_1) \frac{\exp[(T - T_\tau)/s_\tau]}{1 + \exp[(T - T_\tau)/s_\tau]}.$$

모형에서 R 은 Catch Ratio이고, $(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau)$ 는 모형의 모수벡터이다. θ 와 β 는 형상모수, τ_1 은 $T \ll T_\tau$ 일때의 좌극한, τ_2 은 $T \gg T_\tau$ 일때의 우극한, T_τ 는 두 극한사이의 변동에 대한 위치모수이며, s_τ 는 변동을 흐리게 해주는 척도모수이다. s_τ 가 작은 값이라면 빠른 변화를 의미하고, 큰 값이라면 선형모형에 가까워진다 (Wolff et al., 2014). 함수 R 은 일반적으로 풍속에 대해 제 1사분면에 단조감소함수로 나타낼 수 있다.

시간 t 의 단위는 한시간이므로 r_t 는 시간 $t = 1, \dots, T$ 에서 누적 강수량을 이용하여 구해진 Catch Ratio를 의미한다. R_t 는 r_t 에 상응하는 Catch Ratio의 근본적인 참값이다. 각 시간($t=1, \dots, T$)에 대해서, 다음을 따르는 간단한 통계적 모형을 가정할 수 있다.

$$r_t = R_t + \epsilon_t, \epsilon_t \sim N(0, \sigma^2) \text{ or } r_t \sim N(R_t, \sigma^2).$$

여기서, σ^2 는 불확실성의 정도를 설명하는 분산이다. 게다가, 관측값인 $r_1, r_2, r_3, \dots, r_t$ 는 $R_1, R_2, R_3, \dots, R_t$ 가 주어졌을 때 조건부 독립이라고 가정할 수 있다. R_t 는 시간 t 에서 측정된 온도 T 와 풍속 v 그리고 알려지지 않은 모수인 $(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau)$ 의 함수이며 다음과 같이 표현할 수 있다.

$$R_t = R(T, v; \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau).$$

통계적 추론에서 모형의 모수들은 우도를 통해서 추정할 수 있다. 우도함수는 통계적 추론의 현대적 방법에서 가장 기본적인 방법이다. 모수벡터 θ 에 대한 결합밀도함수 $f(y; \theta)$ 에서 데이터 벡터 y 는 Y 의 참값으로 간주된다고 가정하자. 그러면 우도함수는 $\theta \in \Theta$ 에 대해 $L(\theta|y) \propto f(y; \theta)$ 로 정의되고, 우도함수를 이용하여 y 를 기반으로 θ 의 값에 대해 추론할 수 있다. 우도함수는 관심모수를 강조하기 위한 방식으로 표현되고 관측된 데이터는 그것들에 대한 정보를 포함한다. 우도에 근거한 추론은 점추정을 할 수 있고 뿐만 아니라 신뢰영역을 통해서 불확실성을 평가 할 수 있다.

관측된 Catch Ratio의 조건부 독립과 정규성 조건하에서 $r_1, r_2, r_3, \dots, r_t$ 의 우도함수를 다음과 같이 구축 할 수 있다.

$$L(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2 | r_1, r_2, r_3, \dots, r_t) \propto \frac{1}{\sqrt{\sigma^{2T}}} \exp \left\{ - \sum_{t=1}^T \frac{(r_t - R(T_t, v_t; \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau))^2}{2\sigma^2} \right\}.$$

하지만 불행히도, 우도함수에 근거한 추론이 항상 실현 가능한 것은 아니다. 그것은 모형의 구조가 너무 복잡해서 결합밀도함수 $f(y; \theta)$ 에 근거한 추론을 할 수 없거나 통계적 모형의 완전 확률 구조가 구체적으로 명시되지 않기 때문이며, 이는 추론이 불가능하거나 바람직하지 않을 수 있기 때문이다. 따라서 다른 접근법을 이용하여 추론해야 하는 경우가 생긴다. 계산 능력이 발전함으로써, 최근의 논문에서는 복잡한 모형을 평가하기 위해 베이지안 접근법을 이용하고 있으며, 이는 실용적이며 이론적인 장점을 보여 준다.

4. 베이저안 분석

4.1. 사후분포 추정

베이저안 방법은 모형에 포함된 알려지지 않은 모수들에 대한 불확실성의 정도를 제공한다. 게다가, 다모수함수는 수많은 타당할 것 같은 모형의 형태에 대한 가능성을 열어놓는다. 그리고 사전정보를 이 용한다는 측면에서 베이저안 방법은 매우 매력적이다. 계층적 베이저안 모형은 복잡한 문제에 대한 분석에 매우 적합하다는 것을 보여 왔다. 이전의 가정을 토대로, Catch Ratio를 나타내는 다음의 모형을 고려하였다.

$$\begin{aligned} r_t | \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2 &\sim N(R(T_t, v_t; \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau), \sigma^2) \text{ for } t = 1, \dots, T. \\ \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau &\sim \pi(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau). \\ \sigma^2 &\sim \pi(\sigma^2). \end{aligned}$$

모수 집합에 대한 전반적인 분포는 각각의 모수에 대해서 서로 독립을 가정함으로써 구성되고 다음과 같이 표현이 가능하다.

$$\pi(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau) = \pi(\theta) \cdot \pi(\beta) \cdot \pi(\tau_1) \cdot \pi(\tau_2) \cdot \pi(T_\tau) \cdot \pi(s_\tau).$$

여기서는 Catch Ratio에 대하여 시간 t 에서 동일한 분산을 가정하였다. 하지만, 시간 t 에서 동일하지 않은 분산을 가정하여 모형을 확장 할 수 있으며 이를 위해서는 다음과 같은 분산의 유연성 있는 확대가 필요하다. 분산을 다음과 같이 온도에 영향을 받는 함수로 표현하거나,

$$\sigma_t = \sigma_1 + (\sigma_2 - \sigma_1) \frac{\exp[(T_t - T_\tau)/s_\tau]}{1 + \exp[(T_t - T_\tau)/s_\tau]}.$$

또는 시간 t 에서의 온도와 풍속의 일반화선형모형으로 표현 할 수 있다.

$$\log(\sigma_t^2) = \gamma_0 + \gamma_1 T_t + \gamma_2 v_t + e_t, e_t \sim N(0, \delta^2).$$

그렇다면 불확실성을 모형화하기 위해 다음과 같이 모형에 한 단계를 추가하여 확장할 수 있다. 하지만 이번 연구에서는 불확실성의 정도가 시간 t 에 영향을 받지 않는다고 가정하였다.

$$\begin{aligned} r_t | \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma_t^2 &\sim N(R(T_t, v_t; \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau), \sigma_t^2) \text{ for } t = 1, \dots, T. \\ \theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau &\sim \pi(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau). \\ \sigma_t^2 | \gamma_0, \gamma_1, \gamma_2, \delta^2 &\sim \log N(\gamma_0 + \gamma_1 T_t + \gamma_2 v_t, \delta^2). \\ \delta^2 &\sim \pi(\delta^2). \end{aligned}$$

베이저안 방법은 모형의 모수를 평가하기 위해 사용된다. 모수의 사후분포는 모수에 관한 정보를 표현한다. 사후분포들은 사전분포와 우도로부터 마코프 연쇄-몬테칼로 (Markov Chain Monte Carlo ; MCMC) 방법을 이용하여 수치적으로 계산된다. 그 알고리즘은 burn in과 parallel tempering 접근법과 함께 비교적 일반적인 확률보행 메트로폴리스 알고리즘에 근거한다 (Chib과 Greenberg, 2011). 베이저안 데이터 분석과 방법에 대한 세부사항은 Gelman 등 (2013)에 잘 설명되어있다. 앞에서의 모형에 대한 설정을 토대로, 관심모수인 $(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2)$ 에 대한 사후분포를 계산하였고 MCMC를 이용하여 얻어질 수 있다. 추정하고자 하는 모수에 대한 결합사후분포는 다음과 같다.

$$\pi(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2 | r_1, r_2, r_3, \dots, r_t)$$

$$\begin{aligned}
&= \frac{L(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2 | r_1, r_2, r_3, \dots, r_t) \pi(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau)}{m(r_1, r_2, r_3, \dots, r_t)} \\
&\propto \frac{1}{\sqrt{\sigma^{2T}}} \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \\
&\quad \times \pi(\theta) \times \pi(\beta) \times \pi(\tau_1) \times \pi(\tau_2) \times \pi(T_\tau) \times \pi(s_\tau) \times \pi(\sigma^2).
\end{aligned}$$

여기에서 $\tau(T_t) = \tau_1 + (\tau_2 - \tau_1) \frac{\exp[(T_t - T_\tau)/s_\tau]}{1 + \exp[(T_t - T_\tau)/s_\tau]}$ 이고 $m(r_1, r_2, r_3, \dots, r_t)$ 는 관측된 Catch Ratio의 한계밀도함수이며, 이는 결합우도함수에서 모든 모수를 적분하여 얻을 수 있다.

모수 $(\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau, \sigma^2)$ 에 대해서는 서로 독립인 무정보 사전분포를 고려하였다. 추정하고자 하는 모수에 대한 결합사후분포가 일반적으로 알려지지 않은 분포족에 해당함을 알 수 있다. 그러므로 깃스 샘플링을 위해 추정하고자 하는 각 모수의 조건부사후분포를 다음과 같이 나타내었다.

- θ 의 조건부사후분포

$$\theta | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(\theta).$$

$$\text{여기에서 } \tau(T_t) = \tau_1 + (\tau_2 - \tau_1) \frac{\exp[(T - T_\tau)/s_\tau]}{1 + \exp[(T - T_\tau)/s_\tau]}.$$

- β 의 조건부사후분포

$$\beta | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(\beta).$$

- τ_1 의 조건부사후분포

$$\tau_1 | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(\tau_1).$$

- τ_2 의 조건부사후분포

$$\tau_2 | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(\tau_2).$$

- T_τ 의 조건부사후분포

$$T_\tau | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(T_\tau).$$

- s_τ 의 조건부사후분포

$$s_\tau | rest \propto \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(s_\tau).$$

- σ^2 의 조건부사후분포

$$\sigma^2 | rest \propto \frac{1}{\sqrt{\sigma^{2T}}} \exp \left\{ - \sum_{t=1}^T \frac{1}{2\sigma^2} [r_t - [1 - \tau(T_t) \exp[-(v_t/\theta)^\beta] + \tau(T_t)]]^2 \right\} \times \pi(\sigma^2).$$

관심모수에 대한 조건부사후분포 또한 일반적으로 잘 알려지지 않은 분포 족에 해당하는 것을 알 수 있다. 따라서 표본을 생성하기 위해서 메트로폴리스-헤스팅스를 사용하였고, 메트로폴리스-헤스팅스를 위한 후보 생성 밀도함수는 일반적으로 사용되는 확률 보행 연쇄 $q(y|x) = q_1(|y - x|)$ 를 이용하였다.

4.2. 강수자료 적용

예비 탐색적 연구에서, 제안된 모형의 모수는 강수 형태에 따라 동일하지 않음을 보여주었으므로, R 의 특성은 온도에 따라 달라진다는 것으로 가정되었다. 이러한 것은 강수의 형태에 따라, 모형 인덱스 P 를 이용하여 서로 다른 모수의 값을 허용하도록 한다. 즉,

$$R(T, v; \theta^{(P)}, \beta^{(P)}, \tau_1^{(P)}, \tau_2^{(P)}, T_\tau^{(P)}, s_\tau^{(P)}).$$

$$\text{Model index } P = \begin{cases} s, & T \leq -2.0^\circ\text{C}, \\ m, & -2.0^\circ\text{C} < T \leq 3.0^\circ\text{C}, \\ r, & 3.0^\circ\text{C} < T. \end{cases}$$

앞에서 설명한 목표 사후분포로부터 표본을 생성하기 위하여 MCMC방법을 사용하였다. 표본 추출의 효율성과 수렴도를 확인하기 위해 병렬적 구조를 가진 시뮬레이션 체인을 형성하였고, 메트로폴리스-헤스팅스 알고리즘의 기각률과 겔만-루빈통계량 (Gelman-Rubin statistic, G-R statistics)을 각각 확인하였다. 100,000개의 표본을 생성한 다음, 초기값의 영향을 제거하기 위해 처음의 표본 50,000개를 burn in으로 제거하였고, 다음 50,000개의 표본에서부터 자기상관을 없애주기 위해서 매 10번째 MCMC 표본을 선택하였다.

Table 4.1에서는 기온에 의해 구분되어지는 세 가지의 강수형태 (snow, fixed, rain) 하에서 모수 ($\theta, \beta, \tau_1, \tau_2, T_\tau, s_\tau$)들의 사후분포의 평균과 표준오차를 보여준다. 세 가지의 강수형태에서 각 모수들의 분포가 달라지는 것을 확인할 수가 있다. 또한 각 강수형태에 따른 모형의 불확실성을 나타내는 σ 에 대한 사후분포 추정값도 알아보았다. σ 의 사후분포는 강수의 형태에 따라 비슷하게 나타났지만, 값에 대해서는 차이가 나타났다.

Figure 4.1에서는 각 강수형태에 따른 단순선형회귀를 이용하여 적합된 Catch Ratio와 잔차를 보여 준다. 시그모이드 모형을 이용한 적합된 Catch Ratio와 잔차는 Figure 4.2를 통해서 볼 수 있다. 단순 선형회귀는 가장 기본적인 분석방법이며, 기상 관련분야에서는 많이 사용하고 있는 분석방법이기에 시그모이드 모형과 비교해 보았다. 잔차에 대한 MSE의 경우, 단순선형회귀는 0.0294, 시그모이드 모형은 0.0265로 나타났다. 즉, Catch Ratio를 모형에 적합시키기 위해서는 선형모형보다 시그모이드 모형이 좋은 결과가 나타났음을 알 수 있다.

Table 4.1 Estimated coefficients (Coef.) and standard error (SE) values for all components of the proposed models

Precipitation Type Parameter	Rain		Mixed		Snow	
	Est.	SE	EST	SE	EST	SE
θ	2.82	0.981	3.69	1.172	4.52	0.912
β	1.94	0.634	1.69	0.624	2.06	0.455
τ_1	0.14	0.062	0.28	0.093	0.33	0.121
τ_2	0.95	0.013	0.93	0.035	0.94	0.036
T_τ	0.22	0.071	0.49	0.214	0.65	0.24
s_τ	0.21	0.032	1.23	0.297	1.23	0.32
σ	0.03	0.002	0.05	0.002	0.02	0.001

5. 결론

일반적으로 우량계로 측정된 강수량은 지상에 도달한 실제 강수량보다 적게 관측된다. 이는 주로 바람에 의한 강수량의 손실을 발생시키는 계통오차 때문으로 알려져 있는데, 특히 이러한 경향은 고체

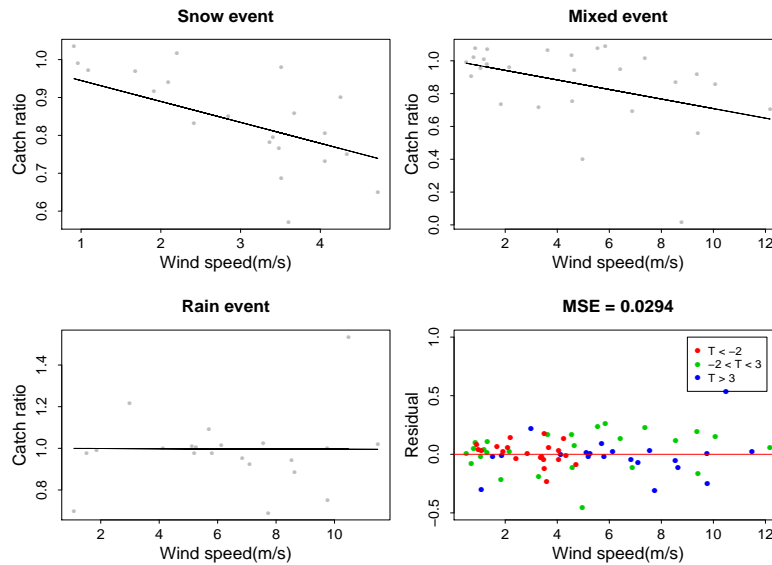


Figure 4.1 Plot of fitted Catch Ratio for each temperature (linear model) and plot of residuals

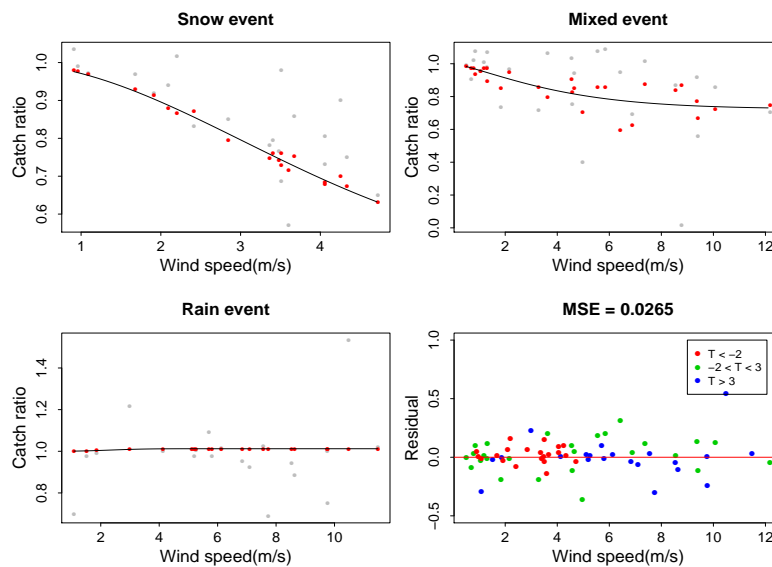


Figure 4.2 Plot of fitted Catch Ratio for each temperature (Sigmoid model) and plot of residuals

강수량의 경우 더 심각하게 발생한다. 본 연구에서는 바람에 의한 고체 강수량의 언더캐치 (under-catch)를 알아보고, 겨울에 내리는 모든 강수의 형태 (snow, mixed, rain)에 대하여 연속조정함수를 적용하여 분석하였다. 이를 위해 고창 표준기상관측소에서 측정된 데이터를 사용하였고, 객관적으로 데이터를 가장 잘 설명하는 모형을 선택하고 평가하기 위해 베이지안 분석을 이용할 것이다. 또한 기상분야에서 일반적으로 사용되는 단순회귀모형과의 비교를 통해 제안된 모형의 적합성을 확인하였다.

제안된 모형의 불확실성에 대하여 다음과 같은 조금 더 일반화된 모형을 가정할 수 있다.

$$\sigma_t = \sigma_1 + (\sigma_2 - \sigma_1) \frac{\exp[(T - T_\tau)/s_\tau]}{1 + \exp[(T - T_\tau)/s_\tau]},$$

또는

$$\log(\sigma_t^2) = \gamma_0 + \gamma_1 T_t + \gamma_2 v_t + e_t, e_t \sim N(0, \delta^2).$$

또한, 이번 결과를 바탕으로 모든 강수형태에 공통으로 적용될 수 있는 모형과 강우의 낙하속도를 함께 고려한 모형으로의 일반화가 가능할 것으로 보인다.

References

- Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, **49**, 327-335.
- Gelman, A., Carlin, B. P., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013). *Bayesian data analysis*, 3rd ed., CRC Press.
- Goodison, B. E., Louie, P. Y. T. and Yang, D. (1998). WMO solid precipitation measurement intercomparison. *WMO Instruments and Observing Methods Rep.67, WMO/TD-No.872*, 212.
- Kim, Y. and Kim, H. J. (2012). Stochastic precipitation modeling based on Korean historical data. *Journal of the Korean Data & Information Science Society*, **23**, 1309-1317.
- Lee, J. E. (2014) *Evaluation of the accuracy of solid precipitation measurements*, Thesis, Kyungpook National University, Daegu.
- Lee, J. J. and Kim, Y. (2016). A spatial analysis of Neyman-Scott rectangular pulse model. *Journal of the Korean Data & Information Science Society*, **27**, 1119-1131.
- Nitu, R. (2013). Cold as SPICE, *Meteorological Technology International*, 148-150.
- Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Thériault, J., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Cristanelli, S., and Gutmann, E. (2012). How well are we measuring snow?. *Bulletin of the American Meteorological Society*, **93**, 811-829.
- Strangeways, I. (2004). Improving precipitation measurement. *International Journal of Climatology*, **24**, 1443-1460.
- Thériault, J. M., Rasmussen, R., Ikeda, K. and Landolt, S. (2012). Dependence of snow gauge collection efficiency on snowflake characteristics. *Journal of Meteorology and Climatology*, **51**, 745-762.
- WMO (2008). *Guide to meteorological instruments and methods of observation*, 7th edition, WMO No. 8, Geneva.
- Wolff, M.A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T. and Brækkan, R. (2015). Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation, results of a Norwegian field study. *Hydrology and Earth System Sciences*, **19**, 951-967.

Bayesian analysis of adjustment function for wind-induced loss of precipitation[†]

Yeongwoo Park¹ · Young Min Kim² · Yongku Kim³

¹²³Department of Statistics, Kyungpook National University

Received 2 March 2017, revised 9 May 2017, accepted 13 May 2017

Abstract

Precipitation is one of key components in hydrological modeling and water balance studies. A comprehensive, optimized and sustainable water balance monitoring requires the availability of accurate precipitation data. The amount of precipitation measured in a gauge is less than the actual precipitation reaching the ground. The objective of this study is to determine the wind-induced under-catch of solid precipitation and develop a continuous adjustment function for measurements of all types of winter precipitation (from rain to dry snow), which can be used for operational measurements based on data available at standard automatic weather stations. This study provides Bayesian analysis for the systematic structure of catch ratio in precipitation measurement.

Keywords: Adjust function, catch ratio, precipitation, wind-induced loss.

[†] This paper is based on part of Yeongwoo Park's Master thesis.

¹ Graduate student, Department of Statistics, Kyungpook National University, Daegu 41566, Korea.

² Assistant professor, Department of Statistics, Kyungpook National University, Daegu 41566, Korea.

³ Corresponding author: Associate professor, Department of Statistics, Kyungpook National University, Daegu 41566, Korea. E-mail: kim.1252@knu.ac.kr