

# A Modeling of Daily Temperature in Seoul using GLM Weather Generator

Hyeonjeong Kim<sup>a</sup> · Hae Young Do<sup>b</sup> · Yongku Kim<sup>a,1</sup>

<sup>a</sup>Department of Statistics, Yeungnam University

<sup>b</sup>Department of Statistics, Kyungpook National University

(Received February 15, 2013; Revised March 27, 2013; Accepted April 26, 2013)

---

## Abstract

Stochastic weather generator is a commonly used tool to simulate daily weather time series. Recently, a generalized linear model(GLM) has been proposed as a convenient approach to fitting these weather generators. In the present paper, a stochastic weather generator is considered to model the time series of daily temperatures for Seoul South Korea. As a covariate, precipitation occurrence is introduced to a relate short-term predictor to short-term predictands. One of the limitations of stochastic weather generators is a marked tendency to underestimate the observed interannual variance of monthly, seasonal, or annual total precipitation. To reduce this phenomenon, we incorporate a time series of seasonal mean temperatures in the GLM weather generator as a covariate.

Keywords: Daily temperature, generalized linear model, overdispersion, stochastic weather generator.

---

## 1. 서론

Stochastic weather generator는 큰 규모의 기상이나 기후정보를 상대적으로 작은 규모의 날씨정보로 downscaling하는 통계기법 중 하나이며, 주로 일일 평균기온이나 일일 강수량을 주어진 기상정보를 활용하여 생성해 내는 통계적인 모형이다 (Wilks와 Wilby, 1999). 실제 이 모형을 통해 특정한 날의 일일 평균기온이나 일일 누적 강수량을 예측하는 것은 거의 불가능하다. 따라서 이 모형은 생성된 일일 날씨를 이용한 계절별 특성치 통계량이 실제 관측된 계절별 특성치 통계량과 유사하도록 구성된다 (Maraun 등, 2010; Wilks, 2010).

최근에는 일반화선형모형(Generalized linear model; GLM)을 이용한 Stochastic weather generator가 많이 사용되는데, 이 모형은 기후의 계절적 특성을 설명하거나 엘니뇨와 같은 현상을 일일 날씨와 연계하여 설명하는데 매우 용의하다 (McCullagh와 Nelder, 1989; Furrer와 Katz, 2007). 하지만 이러한 기법의 한가지 문제점은 모형을 통해 얻어진 계절별 특성값들의 분산이 실제 관측된 계절별 특성값들의 분산보다 상대적으로 작게 추정되는 문제점이 있다 (Buishand, 1978; Katz와 Parlange, 1998). 이러한

---

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (No. 2012-011113).

<sup>1</sup>Corresponding author: Assistant professor, Department of Statistics, Yeungnam University, 280 Daehak-Ro, Gyeongbuk 712-749, Korea. E-mail: ykkim@yu.ac.kr

현상을 “overdispersion” 이라고 하는데, GLM을 이용한 weather generator가 기후의 변동을 충분히 설명할 수 없는 것에 대한 문제를 해결한 것이 상당히 중요한 이슈가 되어왔다 (Katz와 Zheng, 1999).

이러한 overdispersion 현상을 줄이기 위해서, Kim 등 (2012)는 GLM weather generator에 계절별 총 강우량과 계절별 평균 기온을 공변수로 사용하였는데, 직접적인 계절별 통계량의 사용으로 인해 발생할 수 있는 underdispersion 현상을 막기 위해서 smoothing 처리된 계절별 통계량을 모형에 대신 사용하였다. 여기에서는 LOESS(locally weighted scatterplot smoothing; Cleveland, 1979; Hastie와 Tibshirani, 1990) 기법을 이용하여 계절별 통계량을 smoothing 처리하였다. 참고로 다른 smoothing 기법도 고려될 수 있다. Wilks (1989)는 일일 강우량에 관한 모형을 위해서 월별 강우량을 이용하였고, Hansen와 Mavromatis (2001)는 overdispersion 현상을 줄이기 위해서 GLM weather generator의 계수를 임의로 조정하였다. 본 논문에서는 서울지역의 일일 평균기온을 모형화 하기 위해 GLM weather generator를 이용하였다. 모형 적합을 위해서 1961년부터 2011년까지 51년 동안의 서울지역 일일 평균기온 자료가 사용되었다. overdispersion 현상을 줄이기 위해 앞서 소개한 방법과 유사한 기법이 사용되었다. 본 논문은 다음과 같이 구성된다.

2장에서는 일일 평균기온에 관한 GLM weather generator를 간단히 소개하였다. 그리고 일일 평균기온을 모형화하기 위해서 그날의 강우 발생유무가 공변수로 사용되므로 강우 발생을 위한 GLM weather generator도 함께 설명된다. 3장에서는 서울지역의 강우 발생과 일일 평균기온을 GLM weather generator에 적합하였고 실제 자료가 모형에 잘 적합되었는지를 여러 가지 특성값을 이용하여 평가하였다. 그리고 4장에서는 관련된 내용에 대한 추가 논의가 이루어 진다.

## 2. 일일 평균기온에 관한 GLM weather generator

Furrer와 Katz (2007)에 의해서 소개된 일반화 선형모형에 기초한 stochastic weather generator는 Richardson (1981)이 처음으로 제안한 모형 중 가장 간단한 형태이다. 본 논문에서는 이 모형을 간단히 소개한다. 더 자세한 내용은 Furrer와 Katz (2007)나 웹사이트 <http://www.image.ucar.edu/~eva/GLMwgen/>를 참조하면 된다. Furrer와 Katz (2007)에 의해 소개된 강우 발생과 강우 강도에 관한 GLM weather generator는 궁극적으로는 Stern과 Coe (1984)의 모형과 같은 것인데 그들은 모형의 모수가 일년 주기를 가지는 chain-dependent process를 이용하여 일일 강우량에 관한 GLM 모형을 만들었다.

앞에서 잠깐 언급한 바와 같이, downscaling 기법은 영향평가에 사용되는 보다 현실적인 날씨를 생성하기 위한 효과적인 방법 중 하나이다. 규모가 큰 기상현상을 어떻게 규모가 작은 날씨와 연관시키는지에 관한 의문이 서로 다른 두 현상간의 관계를 이해하는 기회를 제공한다. 예를 들어, downscaling 기법은 기후나 대양 같은 큰 규모의 변수와 관측소 위주의 날씨자료 변수의 관계를 주로 이용하게 된다 (Benestad 등, 2008; Katz와 Parlange, 1998; Mehrotra 등, 2004). 특히, 통계적 downscaling은 큰 규모의 기상이나 기후현상과 작은 규모의 날씨사이에 밀접한 관계가 있다는 가정에서 시작한다. 본 논문에서는 제안된 모형을 통해서 downscaling 기법을 GLM weather generator에 자연스럽게 도입하고자 한다.

Kim 등 (2012)은 LOESS smoothing된 계절별 기상 통계치를 기존의 GLM weather generator에 추가 변수로 고려함으로써 일일 날씨 통계량에 충분한 불확실성을 소개하여 “underdispersion” 현상을 제거하였다. 이 접근법은 큰 규모의 설명변수를 작은 규모의 종속변수에 연결하는 것에 중점을 두고 있는데, 예를 들어, 큰 규모의 기후나 대류 현상 등을 나타내는 여러 가지 지수가 일일 강우모형에서 설명변수로 사용된다. 하지만 이 GLM weather generator에서는 주어진 큰 규모의 특성값(설명변수)을 작은 규모

의 여러 값(중속변수)으로 분해(disaggregation)하는 역할을 한다. 즉, 계절별 총 강우량이나 계절별 평균 온도를 이용하여 일일 강우량이나 일일 기온을 생성하는 모형을 만든다.

GLM weather generator에 smoothing되지 않은 계절별 기후 특성값을 직접 공변수로 사용할 경우 일일 날씨모형에 오히려 너무 과한 변동(즉, underdispersion 현상)이 나타날 수 있으므로 우리는 적당한 수준의 smoothing된 계절별 기후 특성값을 공변수로 사용한다. 이 때 사용될 계절별 기후 특성값을 smoothing하기 위해서 LOESS (Cleveland, 1979)기법을 사용하였다. LOESS는 선형회귀모형에서 최소제곱법이 가지는 간단함과 비선형회귀모형의 유연성을 적절히 결합한 방법으로, 간단히 locally weighted polynomial regression으로 설명되기도 한다. 이 방법은 간단하면서도 매우 유연하게 사용할 수 장점이 있으며 이상값에 의한 영향도 상대적으로 적어 Robust한 smoothing 결과를 제공한다. 특히, LOESS 방법은 특정한 함수를 지정할 필요없이 smoothing 모수만으로 조절이 가능하다. 참고로 본 논문에서는 사전 분석을 통하여 smoothing 모수 0.3으로 하였다. 결론적으로, 우리는 서울지역의 일일 기온을 모형화하기 위해서 LOESS smoothing된 계절별 특성값(평균기온)을 기존의 GLM weather generator에 공변수로 추가하였다.

일반적으로 어느 날의 기온은 그 날의 강우발생 유무에 따라 영향을 받을 수 있으므로, 일일 기온에 대한 GLM weather generate에 강우발생 유무를 공변수로 고려할 필요가 있다. 따라서 일일 기온에 대한 모형화에 앞서 강우발생에 대한 모형을 정의하고자 한다. 먼저  $J_t$ 를 어떤 특정한  $t$ 날의 강우발생 상태를 나타내는 변수라고 하자 (즉,  $t$ 날 비가 오면  $J_t = 1$ 이고 그렇지 않으면  $J_t = 0$ 이다). 이 때,  $t$ 날 비가 올 확률을  $p_t$ 라고 정의하자 (즉,  $p_t = P\{J_t = 1\}$ 이고  $1 - p_t = P\{J_t = 0\}$ 이 된다). 그리고 두 개의 상태를 가지는 일차 Markov chain에서처럼, 그 날의 비가 올 확률은 그 전날 비가 왔는지 오지 않았는지에 따라 달라질 수 있다. 그러므로  $t$ 날의 강우발생에 대한 모형은 다음과 같이 표현할 수 있다.

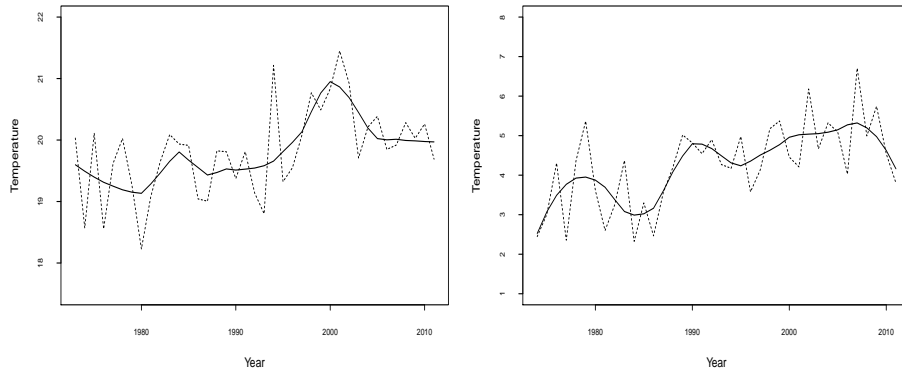
$$\ln\left(\frac{p_t}{1-p_t}\right) = \mu + \alpha J_{t-1} + \beta_1 C_t + \beta_2 S_t + \gamma_1 C_t J_{t-1} + \gamma_2 S_t J_{t-1} + \beta_s I_t P_t^S + \beta_w (1 - I_t) P_t^W, \quad (2.1)$$

여기에서  $I_t$ 는 계절을 나타내는 계절지수이다 (즉, 우기(4월에서 9월까지)일 때는  $I_t = 1$ 이고 건기(10월에서 3월까지)일 때는  $I_t = 0$ 이 된다). 또 일년 주기의 특성을 설명하기 위해서  $C_t = \cos(2\pi(t - 181)/365)$ 와  $S_t = \sin(2\pi(t - 181)/365)$ 를 모형에 추가하였고,  $P_t^S$ 와  $P_t^W$ 는 각각 LOESS smoothing된 우기와 건기의 계절별 누적 강우량이다. 참고로, 이 값들은 우기와 건기에 따라 구분해서 각각 smoothing되었다. 여기에서 모형 계수  $\alpha$ 는 그 전날 강우발생 유무에 따라 그 날의 강우발생 확률이 어떻게 달라지는지를 나타내고,  $\beta_1$ 과  $\beta_2$ 는 강우가 발생할 확률에 대한 주기를 나타내는 SINE 곡선의 위상과 진폭을 결정한다.  $\gamma_1$ 와  $\gamma_2$ 는 이러한 주기 효과를 그 전날의 강우 발생유무에 영향을 받도록 허용해 준다.

이제 일일 기온에 대한 GLM weather generate를 알아보자. 먼저  $X_t$ 를 어떤 특정한  $t$ 날의 일일 기온이라고 하자. 그 날의 기온은 그 전날의 기온에 의존하는 경우가 많으므로 Richardson (1981)처럼 first-order autoregressive [AR(1)] 모형을 고려한다.

$$X_t = \mu_X + \alpha_X J_t + \varphi X_{t-1} + \beta_{X,1} C_t + \beta_{X,2} S_t + \beta_{X,S} I_t T_t^S + \beta_{X,W} (1 - I_t) T_t^W + \varepsilon_{X,t}, \quad (2.2)$$

여기에서  $T_t^S$ 와  $T_t^W$ 는 LOESS smoothing된 우기와 건기의 각각 평균 기온이다. 그리고 계절을 나타내는 변수  $I_t$ 는 각 계절에 따라 기후/기상 조건과 일일 기온간의 관계가 다를 수 있도록 허용한다. 오차항인  $\varepsilon_{Y,t}$ 는 평균이 0이고 분산  $\sigma^2$ 인 정규분포를 따른다. 모형계수  $\alpha_X$ 는 그날의 강우발생 유무에 따라 평균기온이 어떻게 달라지는 지를 보여주고,  $\beta_{X,1}$ 와  $\beta_{X,2}$ 는 강우모형에서처럼 일년주기의 SINE 곡선을 결정한다. 그리고  $\varphi$ 는 일일 기온의 자기상관 정도를 나타낸다. 마지막으로  $\beta_{X,S}$ 와  $\beta_{X,W}$ 는 계절별(우기/건기) 평균기온이 일일 기온에 미치는 영향을 나타낸다. 참고로, 분석의 간편성을 위해서



**Figure 2.1.** Optimal smoothed aggregated covariates of seasonal mean temperature during summer (left) and winter (right) seasons. Dashed lines: corresponding observed values of the data series.

**Table 3.1.** Estimated coefficients(Coef.) and standard error(SE) values for all components of the stochastic weather generators with aggregated climate statistics as covariates at Seoul.

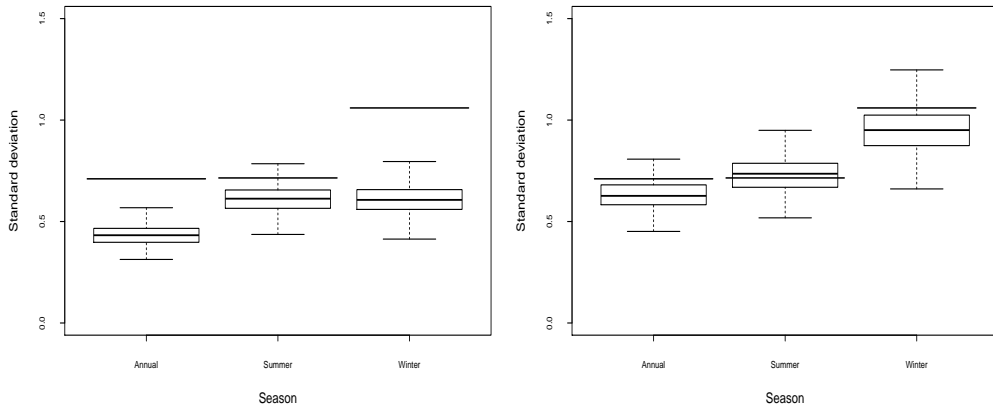
Covariate category	Precipitation occurrence			Daily mean temperature		
	Term	Coef.	SE	Term	Coef.	SE
Constant	$\mu$	-1.70	0.117	$\mu$	2.31	0.124
Precipitation	-	-	-	$J_t$	-0.52	0.039
Wet	$I_t P_t^S$	0.07	0.022	$I_t T_t^S$	0.048	0.007
Dry	$(1 - I_t) P_t^W$	0.33	0.116	$(1 - I_t) T_t^W$	0.21	0.025
Autocorrelation	$J_{t-1}$	0.98	0.042	$X_{t-1}$	0.74	0.006
Seasonality	$C_t$	-0.17	0.064	$C_t$	-3.19	0.087
	$S_t$	-0.08	0.034	$S_t$	-1.23	0.039
Interaction	$C_t J_{t-1}$	-0.47	0.058	-	-	-
	$S_t J_{t-1}$	-0.15	0.059	-	-	-
AIC	15711.0			60774.9		
BIC	15771.5			60835.4		

강우모형 (2.1)과 일일 기온모형 (2.2) 모두에 smoothing 모수로서 0.3을 사용하였다. 또한 추가 분석에서 smoothing 모수의 변화에 따라 overdispersion 결과가 robust함을 확인하였다. Figure 2.1은 1961년부터 2011년까지 51년 동안의 서울지역의 계절별 평균 기온을 우기와 건기에 따라 각각 LOESS smoothing한 결과를 보여준 것이다.

### 3. 서울지역 일일기온 관측자료에 적용

한국 기상청에서 제공한 1961년부터 2011년까지 51년동안의 서울지역 일일 기온 자료를 제안된 GLM weather generator에 적합 시켜보았다. 기온 자료의 연간 주기를 보았을 때 여름에 최고가 되고 겨울에 최저가 됨을 명확하게 확인할 수 있었다. 이 기간의 자료들은 거의 결측치 없이 잘 조사되어 있어 모든 자료들이 분석에 사용되었다. 단, 분석과 프로그래밍의 편의를 위해서 2월 29일 자료는 삭제 후 사용하였다. 강우발생에 관한 모형 추정을 위해서 동일 기간의 서울지역 강우발생 자료를 이용하였다.

Table 3.1에서는 기상청에서 제공된 서울지역 자료를 이용하여 강우발생 유무와 일일 기온에 대한 GLM weather generator에 사용된 모수의 추정량과 추정오차를 각각 보여준다. 참고로 분석에는 공개



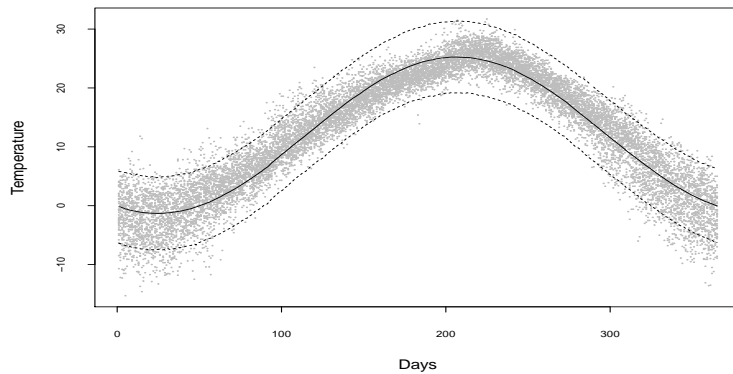
**Figure 3.1.** Boxplots of simulated Sd of annual, summer and winter mean temperature based on the model without aggregated covariates (left) and with smoothed aggregated covariates (right). Horizontal solid lines: corresponding observed values of the data series.

소스 통계 프로그램인 R의 “glm”과 “lm” 함수를 사용하였다. 분석표에서 나타난 바와 같이 사용된 모든 변수들이 통계적으로 유의함을 알 수 있다. 참고로, 강우발생을 공변수로 사용한 일일 기온 모형이 그렇지 않은 모형보다 AIC와 BIC 기준에서 더 좋은 모형으로 확인되었다. 이 때, 다른 변수에 대한 계수 추정량들은 강우발생의 유무에 관계 없이 큰 변화가 없었다.

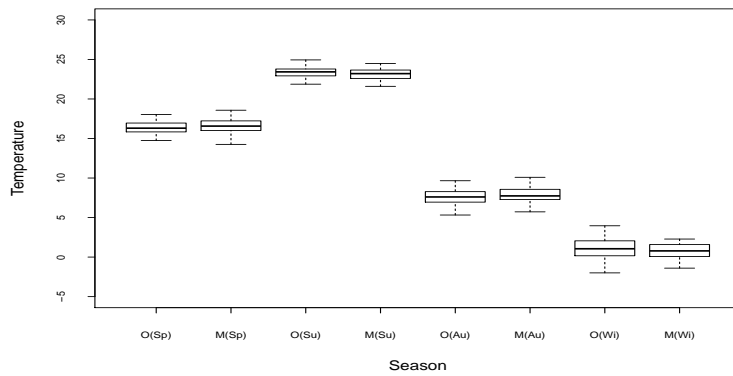
제안된 모형의 유효성을 평가하기 위해 추정된 모형에서 생성된 값들을 이용한 여러 가지 계절별 통계 특성치와 실제 관측된 자료들의 특성값들을 비교하였다. 이를 위해서 실제 관측기간과 동일한 51년치의 일일 기온을 모형으로부터 생성하는데, 이러한 작업을 500회 반복하였다. 참고로 일일 기온 생성에 필요한 일일 강우발생 유무는 GLM weather generator에서 동시에 생성하여 일일 기온모형에 변수로 입력된다.

Figure 3.1은 일일 기온 모형에 LOESS smoothing된 계절별 평균 기온을 사용한 모형이 그렇지 않은 모형보다 얼마나 overdispersion 면에서 향상되었는지를 계절별 또는 연간 평균 기온의 분산을 통해서 비교하였다. 건기의 경우 여전히 약간의 overdispersion 현상이 남아 있는데 이는 상대적으로 건기의 연간 계절별 온도 변화가 우기에 비해 심하지 않아 충분한 변동을 모형에 추가하지 못한 이유로 생각할 수 있다. 하지만 기존의 모형에 비교해서 overdispersion 현상이 개선 되었음을 확인할 수 있었다. 참고로 상자그림은 통계실험을 통한 변동의 범위를 나타내거나 실제 관측된 값들과 직접적으로 비교하는데 효과적인 방법 중 하나로 여러 응용분야에서 많이 사용되고 있다. 앞에서 잠깐 언급한 바와 같이 여기에서 사용된 smoothing 모수값은 결과에 아주 민감하지는 않으므로 적당한 수준에서 결정하여 사용하면 된다. 그리고 smoothing을 적게 하면 할수록 너무 많은 변동이 모형에 전해져서 오히려 underdispersion 현상이 발생할 수 있으므로 적당한 smoothing 모수값을 정하는 것은 반드시 필요하다.

추정된 GLM weather generator를 평가하기 위해서 또 다른 의미 있는 일일 기온에 관련된 통계치를 비교해 보자. Figure 3.2는 모형을 통해 얻은 일일 기온값들을 1년 주기로 하여 각 날의 평균값과 그 95% 신뢰구간을 실제 51년간의 관측값을 함께 나타내 비교하였다. 대체적으로 모형이 실제 일일 기온 관측값들의 1년 주기를 비교적 잘 표현한다고 할 수 있다. 이와 함께, 실제 관측값들의 4계절 평균값과 모형이 생성해낸 자료의 4계절 평균값을 상자그림을 통해 비교하였는데 여기에서도 상당히 유의한 결과를 얻을 수 있었다 (Figure 3.3을 참고).



**Figure 3.2.** 95% confidence intervals (dotted lines) and mean temperatures (solid line) based on model outputs. Dots: observed temperatures on each day of the year.



**Figure 3.3.** Boxplots of observed seasonal mean temperatures (O) and simulated seasonal mean temperatures (M): Sp(March-May), Su(June-August), Au(September-November), Wi(December-February).

#### 4. 결론

본 논문에서 우리는 GLM weather generator를 1961년부터 2011년까지 51년간의 서울지역 일일 기온에 적용하였다. 이를 통해 일일 강우발생이라는 단기 현상과 smoothing된 계절별 평균기온이라는 장기 현상이 일일 기온이라는 단기 현상에 어떻게 영향을 미치는지 알아보았고, 이를 통해 일반적으로 GLM weather generator가 가지는 계절별 평균기온의 overdispersion 현상을 효과적으로 제거할 수 있었다. 그리고 이를 통해 보다 실질적인 일일 기온을 생성할 수 있을 뿐만 아니라 더 효과적인 위험평가도 가능해진다. 물론 LOESS에서 적절한 smoothing 정도를 결정하는 과정이나 계절별 평균기온의 일일 기온 모형에 재사용이라는 관점에서 여전히 문제점을 가지고 있지만 기후/기상 예측이라는 관점에서는 미래의 주어진 시나리오(주로 계절별 평균기온 등)에 따라 일일 날씨가 어떻게 변하는지 또 이를 이용해 계절별 특성값이 어떻게 변하는지를 효과적으로 알아볼 수 있으므로 이러한 접근법은 여전히 유용한 방법이라고 할 수 있다. 서울지역의 일일 기온에 적용된 모형은 실제 관측값과 유사한 계절별 통계 특성치를 보여주는 유용한 결과를 제공하였다.

일반적으로 통계적 downscaling은 보다 큰 규모의 기후/기상 상태나 지역의 지문학적 정보를 이용하여 작은 규모의 지역날씨를 알아보기 위해 사용되는데, 특히 제안된 모형은 긴 기간의 기후를 이용하여 짧

은 기간의 날씨를 직접적으로 추정하는 temporal downscaling을 제공한다. 그리고 한 지역이 아니라 여러 지역에 관한 모형으로 확장하는데도 GLM weather generator는 상대적으로 유용하다. 하지만 여기에 관한 좀 더 다양한 연구가 추가적으로 필요하다.

## References

- Benestad, R. E., Hanssen-Bauer, I. and Chen, D. (2008). *Empirical-Statistical Downscaling*, World Scientific, Singapore.
- Buishand, T. A. (1978). Some remarks on the use of daily rainfall models, *Journal of Hydrology*, **47**, 235–249.
- Cleveland, W. S. (1979). Robust locally-weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, **74**, 829–836.
- Furrer, E. M. and Katz, R. W. (2007). Generalized linear modeling approach to stochastic weather generators, *Climate Research*, **34**, 129–144.
- Hansen, J. W. and Mavromatis, T. (2001). Correcting low-frequency variability bias in stochastic weather generators, *Agricultural and Forest Meteorology*, **109**, 297–310.
- Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Chapman and Hall, New York.
- Katz, R. W. and Parlange, M. B. (1998). Overdispersion phenomenon in stochastic modeling of precipitation, *Journal of Climate*, **11**, 591–601.
- Katz, R. W. and Zheng, X. (1999). Mixture model for overdispersion of precipitation, *Journal of Climate*, **12**, 2528–2537.
- Kim, Y., Katz, R. W., Rajagopalanc, B., Furrer, E. M. and Podestá, G. (2012). Reducing overdispersion in stochastic weather generators using a generalized linear modeling approach, *Climate Research*, **53**, 13–24.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M. and Thiele-Eich, I. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Reviews of Geophysics*, **48**, doi:10.1029/2009RG000314.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed., Chapman and Hall.
- Mehrotra, R., Sharma, A. and Cordery, I. (2004). Comparison of two approaches for downscaling synoptic atmospheric patterns to multisite precipitation occurrence, *Journal of Geophysical Research D: Atmospheres*, **109**, 1–15.
- Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, **17**, 182–190.
- Stern, R. D. and Coe, R. (1984). A model fitting analysis of daily rainfall data, *Journal of the Royal Statistical Society A*, **147**, 1–34.
- Wilks, D. S. (1989). Conditioning stochastic daily precipitation models on total monthly precipitation, *Water Resources Research*, **25**, 1429–1439.
- Wilks, D. S. (2010). Use of stochastic weather generators for precipitation downscaling, *Wiley Interdisciplinary Reviews: Climate Change*, **1**, doi:10.1002/wcc.85
- Wilks, D. S. and Wilby, R. L. (1999). The weather generator game: A review of stochastic weather models, *Progress in Physical Geography*, **23**, 329–357.

# GLM 날씨 발생기를 이용한 서울지역 일일 기온 모형

김현정<sup>a</sup> · 도해영<sup>b</sup> · 김용구<sup>a,1</sup>

<sup>a</sup>영남대학교 통계학과, <sup>b</sup>경북대학교 통계학과

(2013년 2월 15일 접수, 2013년 3월 27일 수정, 2013년 4월 26일 채택)

---

## 요약

확률적 날씨 발생기(Stochastic weather generator)는 일일 날씨를 생성하는데 일반적으로 사용되는 방법으로 최근에는 일반화선형모형에 기초한 확률적 날씨 발생 방법이 제안되었다. 본 논문에서는 서울지역의 일일 기온을 모형화 하하기 위해서 일반화선형모형에 기초한 확률적 날씨 발생기를 고려하였다. 이 모형에서는 계절성을 나타내는 변수와 강우발생 유무가 공변수로 사용되었다. 일반적으로 확률적 날씨 발생기에서는 생성된 일일 날씨가 월별 또는 계절별 총강우량이나 평균온도에 충분한 변동을 만들어 내지 못하는 과대산포 현상이 발생하는데, 이러한 한계를 극복하기 위해 본 연구에서는 평활된 계절별 평균 온도를 일반화선형모형의 공변수로 추가하였다. 그리고 제안된 모형을 1961년부터 2011년까지 51년 동안의 서울지역 일일 평균 기온자료에 적용하였다.

주요용어: 일일기온, 일반화선형모형, 과대산포, 확률적 날씨 발생기.

---

---

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2012-011113).

<sup>1</sup>교신저자: (712-749) 경상북도 경산시 대학로 280, 영남대학교 통계학과, 조교수. E-mail: ykkim@yu.ac.kr